OCT 19 1948

THE QUARTERLY JOURNAL OF
MATHEMATICS

OXFORD SERIES

Volume 19 *No.* 75 *September* 1948

CONTENTS

H. M. Sengupta: On Continuous Independent Functions 129

V. C. Morton and **M. T. Chapple:** A Point Representation of a System of Space Cubic Curves which pass through four given Points and whose Chords belong to a Given Tetrahedral Complex . 133

P. A. P. Moran: A Class of Complex Markoff Chains 140

H. A. Buchdahl: The Hamiltonian Derivatives of a Class of Fundamental Invariants . . . 150

S. Schwarz: On the Equation $a_1 x_1^k + a_2 x_2^k + \ldots + a_k x_k^k + b = 0$ in Finite Fields . . . 160

E. C. Titchmarsh: Complex Fourier-Bessel Transforms 164

H. R. Pitt: A Note on Some Elementary Tauberian Theorems 177

K. L. Chang: On Some Diophantine Equations $y^2 = x^3 + k$ with no Rational Solutions . . 181

H. Davenport and **Marshall Hall:** On the Equation $ax^2 + by^2 + cz^2 = 0$ 189

OXFORD
AT THE CLARENDON PRESS
1948

Price 7s. 6d. net

PRINTED IN GREAT BRITAIN BY CHARLES BATEY AT THE OXFORD UNIVERSITY PRESS

THE QUARTERLY JOURNAL OF MATHEMATICS

OXFORD SERIES

Edited by T. W. CHAUNDY, U. S. HASLAM-JONES, J. H. C. THOMPSON

With the co-operation of A. L. DIXON, W. L. FERRAR, E. A. MILNE, E. C. TITCHMARSH

THE QUARTERLY JOURNAL OF MATHEMATICS (OXFORD SERIES) is published at 7s. 6d. net for a single number with an annual subscription (for four numbers) of 27s. 6d. post free.

Papers, of a length normally not exceeding 20 printed pages of the Journal, are invited on subjects of Pure and Applied Mathematics, and should be addressed 'The Editors, Quarterly Journal of Mathematics, Clarendon Press, Oxford'. The Editors as a rule will not wish to accept material that they cannot see their way to publish within a twelvemonth. While every care is taken of manuscripts submitted for publication, the Publisher and the Editors cannot hold themselves responsible for any loss or damage. Authors are advised to retain a copy of anything they may send for publication. Authors of papers printed in the Quarterly Journal will be entitled to 50 free offprints. Correspondence on the *subject-matter* of the Quarterly Journal should be addressed, as above, to 'The Editors', at the Clarendon Press. All other correspondence should be addressed to the Publisher

GEOFFREY CUMBERLEGE
OXFORD UNIVERSITY PRESS
AMEN HOUSE, LONDON, E.C. 4

BOWES & BOWES

University Booksellers

PAY GOOD PRICES FOR

MATHEMATICAL BOOKS

Offers of Libraries or smaller collections of Mathematical and Scientific Books and Journals, English and Foreign, are invited

1 & 2 TRINITY STREET, CAMBRIDGE

THEORY OF PROBABILITY

By HAROLD JEFFREYS. *Second edition*, 1948. 30*s. net*

(INTERNATIONAL SERIES OF MONOGRAPHS ON PHYSICS)

A new and revised edition of this book the chief object of which is to provide a method of drawing inferences from observational data that will be self-consistent and can be used in practice. It is based on a definite recognition that inductive arguments cannot be brought within the scope of deductive logic, and need additional postulates in any case. They are stated as principles of the method of analysis, not as postulates about the external world; postulates about the external world can be discussed by the theory but do not form part of it. The principal types of problem treated in current statistical theory are discussed in detail and numerous specific applications are given. A new method of stating prior probabilities achieves much greater generality than that given in the first edition.

OXFORD UNIVERSITY PRESS

BLACKWELL'S

University Booksellers

want to buy

Standard books on MATHEMATICS

BROAD STREET, OXFORD

CHAMBERS'S

FOUR-FIGURE MATHEMATICAL TABLES

By L. J. COMRIE, M.A., Ph.D.

A completely new collection by a leading authority. Notable features—all six trigonometrical functions with both decimal and sexagesimal arguments, adequate tabulation of circular, hyperbolic and exponential functions, up-to-date tables of mathematical and physical constants. Throughout means are provided for accurate interpolation.

Some expert opinions:

'This book of tables strikes me as one of the best I have seen. . . .'

'I think these tables are an excellent production—Dr. Comrie's editorship would ensure that—and I shall bring them to the attention of students and others as occasion allows.'

'There is no doubt that such a set of tables was greatly needed, and I find this set admirable in every respect.'

'There are many new features, and I am convinced that the Tables will be found unusually serviceable. One thing stands out immediately—their extraordinary legibility; no strain whatever is imposed on the eye.'

Limp cloth, 5s. *Cloth boards, 6s.*

In preparation

Chambers's SIX-FIGURE MATHEMATICAL TABLES by L. J. Comrie, M.A., Ph.D.

W. & R. CHAMBERS LTD., 38 SOHO SQUARE, LONDON, W.1

HEFFER'S OF CAMBRIDGE

will pay good prices for

MATHEMATICAL BOOKS

especially the following

Boyle, Works; **Newton,** Works; **Cavendish,** Scientific Papers; **Cayley,** Collected Mathematical Papers; **Glazebrook,** Dictionary of Applied Physics; **Faraday,** Experimental Researches in Electricity; **Faraday,** Diaries, 7 vols.; **Gibbs,** Collected Works; **Hilton,** Mathematical Crystallography; **Ramanujan,** Collected Papers; **Rayleigh,** Scientific Papers; **Sylvester,** Collected Scientific Papers

and for standard books on any subject

W. HEFFER & SONS LIMITED
Petty Cury Cambridge

ON CONTINUOUS INDEPENDENT FUNCTIONS

By H. M. SENGUPTA (*Dacca*)

[Received 19 February 1947; in revised form 18 December 1947]

In an illuminating paper entitled 'Note on continuous independent functions', Offord† determines a sufficient condition that two continuous functions defined on $0 \leqslant x \leqslant 1$ may fail to be independent.

The independence of two real functions $f(x)$ and $g(x)$, both given on $0 \leqslant x \leqslant 1$, is defined as follows: Consider all sets

$$E_1 \equiv E_x[\alpha \leqslant f(x) \leqslant \beta] \quad \text{and} \quad E_2 \equiv E_x[\alpha' \leqslant g(x) \leqslant \beta'],$$

where α, β and α', β' are any pairs of real numbers. The functions $f(x)$ and $g(x)$ are said to be *independent* if

$$|E_1 E_2| = |E_1| . |E_2|.$$

In the note referred to above, Offord establishes the theorem

OFFORD'S THEOREM. *Supposing that $f(x)$ and $g(x)$ are both continuous on the closed interval I $(0 \leqslant x \leqslant 1)$ and that neither is a constant, if either of the functions assumes at least one of its values at most enumerably often, then $f(x)$ and $g(x)$ cannot be independent.*

From the above Offord draws the conclusion that, if two continuous functions are independent in the closed interval I and if neither is a constant, then each function takes every one of its values an unenumerable infinity of times. In the present note I point out an extension of the above theorem and I give a proof planned on different lines.

The following preliminaries are necessary before we can enunciate the theorem.

We consider two functions $f(x)$ and $g(x)$ which are continuous in the closed interval I $(0 \leqslant x \leqslant 1)$. Let a be one of the values of $g(x)$ and $E_{g=a} = E_x[g(x) = a]$ the set of points x in $0 \leqslant x \leqslant 1$ for which $g(x)$ has the value a. We define $F_a = f[E_{g=a}]$ as the set of values assumed by $f(x)$ at the points $E_{g=a}$. Since $g(x)$ is continuous, $E_{g=a}$ is a closed non-empty set and, since $f(x)$ is continuous, F_a is a closed set for every value a taken by $g(x)$.

† A. C. Offord, *Quart. J. of Math.* (Oxford), 12 (1941), 86–8.

The extension referred to above may now be given in the form of the theorem:

THEOREM. *If the set F_a is not identical with the entire range of values of $f(x)$ for every value a of $g(x)$, then the functions $f(x)$ and $g(x)$ are not independent.*

In particular, if $F_a \equiv f[E_{g=a}]$ be non-dense, which is certainly the case when $E_{g=a}$ is enumerable, $f(x)$ and $g(x)$ cannot be independent.

Proof. Since $F_a \equiv f[E_{g=a}]$ is not identical with the entire interval $l \leqslant x \leqslant m$, where l and m are the lower and upper bounds of $f(x)$, and $f(x)$ is continuous, there is a value b of $f(x)$ lying in the interval $l \leqslant y \leqslant m$ which does not occur in F_a. The remaining portion of the proof is word for word identical with that given in Offord's paper. The success of the proof depends on the possibility of the selection of the number b lying in $l \leqslant y \leqslant m$, which value is not assumed by $f(x)$ for any x that lies on $E_{g=a}$.

The following is an immediate corollary: *If two continuous functions $f(x)$ and $g(x)$ be independent, then for each set $E_{g=a}$ at which $g(x) = a$, where $L \leqslant a \leqslant M$ and L and M are lower and upper bounds respectively of $g(x)$, the set $F_a = f[E_{g=a}]$ will be identical with the entire interval $l \leqslant y \leqslant m$. Also, if $E_{f=a'}$ be the set at every point of which $f(x) = a'$ $(l \leqslant a' \leqslant m)$, the set $g[E_{f=a'}]$ will be the entire interval $L \leqslant y \leqslant M$.*

In particular, if $E_{g=a}$ be enumerable, then $f[E_{g=a}]$ is necessarily so, and accordingly it cannot consist of the entire interval $l \leqslant y \leqslant m$.

I now give an alternative proof of the theorem.

Alternative proof. Denote by the symbol $E_{g=a}$ the set over which $g(x) = a$. Let $\delta_1, \delta_2, \delta_3, \ldots, \delta_n, \ldots$ be a decreasing sequence of positive numbers such that δ_n tends to zero as n tends to infinity. Let a be a number of the interval $L \leqslant y \leqslant M$ for which we suppose that the set $f[E_{g=a}]$ is not identical with the interval $l \leqslant y \leqslant m$. Let

$$E_x[a-\delta_n \leqslant g(x) \leqslant a+\delta_n]$$

stand for the set over which $a-\delta_n \leqslant g(x) \leqslant a+\delta_n$.

Evidently

$$E_x[a-\delta_1 \leqslant g \leqslant a+\delta_1] \supset E_x[a-\delta_2 \leqslant g \leqslant a+\delta_2] \supset \ldots$$
$$\supset E_x[a-\delta_n \leqslant g \leqslant a+\delta_n] \supset \ldots;$$

and

$$\lim_{n\to\infty} E_x[a-\delta_n \leqslant g \leqslant a+\delta_n] = E_{g=a}.$$

Now, since the set $f[E_{g=a}]$ is not the entire interval $l \leqslant y \leqslant m$ and since it is closed, we can select an interval $\alpha \leqslant y \leqslant \beta$ lying in $l \leqslant y \leqslant m$ such that the set $E_x[\alpha \leqslant f(x) \leqslant \beta]$ for which $\alpha \leqslant f(x) \leqslant \beta$ has no point in common with $E_{g=a}$. Again the sets $E_{g=a}$ and $E_x[\alpha \leqslant f(x) \leqslant \beta]$ are both closed and $E_{g=a}.E_x[\alpha \leqslant f(x) \leqslant \beta] = 0$, so that the distance between the two sets is a fixed positive number, ϵ say.

Enclose each of the points of $E_{g=a}$ in a closed interval of length less than $\frac{1}{2}\epsilon$ with the point as its mid-point. A finite number of these intervals will cover the entire set $E_{g=a}$ so that any point of $E_{g=a}$ is an interior point of at least one of these intervals. Denote the set made up of the sum of these closed intervals by S. Then S is closed and $S.E_x[\alpha \leqslant f(x) \leqslant \beta] = 0$.

Now $E_{g=a} \subset S$, and $E_x[a-\delta_n \leqslant g \leqslant a+\delta_n]$ is a monotonically decreasing sequence of closed sets containing $E_{g=a}$ and tending to $E_{g=\alpha}$ in the limit. We can therefore find an integer N such that

$$E_x[a-\delta_n \leqslant g \leqslant a+\delta_n] \subset S \quad \text{for} \quad n \geqslant N.$$

Choose one such $E_x[a-\delta_n \leqslant g \leqslant a+\delta_n]$ calling it E_1 and denote the set $E_x[\alpha \leqslant f(x) \leqslant \beta]$ by E_2. It is evident that $|E_1| > 0$ and $|E_2| > 0$, so that $|E_1|.|E_2| > 0$, while $E_1.E_2 \subset S.E_2 = 0$. Thus

$$|E_1|.|E_2| > |E_1 E_2| = 0.$$

The functions $f(x)$ and $g(x)$ are therefore not independent.

It is obvious that what we have proved in this theorem is really equivalent to the following:

If $f(x)$ and $g(x)$ be continuous on $0 \leqslant x \leqslant 1$, then a necessary condition for $|E_1 E_2| > 0$ is that the set of values assumed by $f(x)$ over the set of points where $g(x)$ has the value a $(L \leqslant a \leqslant M)$ must be identical with the entire interval $l \leqslant y \leqslant m$ and also that the set of values assumed by $g(x)$ over the set of points where $f(x)$ has the value a', where $l \leqslant a' \leqslant m$, must be identical with the entire interval

$$L \leqslant y \leqslant M.$$

It is interesting to note that the converse of the above is also true. In other words, we can prove that the conditions set out for $f(x)$ and $g(x)$ in the above paragraph are also sufficient that

$$|E_1 E_2| > 0.$$

Proof. Let $E_1 \equiv E_x[\alpha \leqslant f(x) \leqslant \beta]$, where $l \leqslant \alpha < \beta \leqslant m$ and $E_2 \equiv E_x[\alpha' \leqslant g(x) \leqslant \beta']$, where $L \leqslant \alpha' < \beta' \leqslant M$.

Further, let

$$E_1^0 \equiv E_x[\alpha < f(x) < \beta] \quad \text{and} \quad E_2^0 \equiv E_x[\alpha' < g(x) < \beta'].$$

Since $f(x)$ and $g(x)$ are continuous, both E_1^0 and E_2^0 are open sets. Now $g(E_1^0) \supset g[E_{f=\gamma}]$, where γ is a number such that $\alpha < \gamma < \beta$.

But, by hypothesis, $g[E_{f=\gamma}]$ is identical with the entire interval $L \leqslant x \leqslant M$. We denote this interval by the symbol $\langle L, M \rangle$. It follows that

$$g(E_1^0) \supset \langle L, M \rangle.$$

Also

$$g(E_1^0) \subset \langle L, M \rangle.$$

Thus

$$g(E_1^0) \equiv \langle L, M \rangle,$$

and so E_1^0 must contain a sub-set at every point of which

$$\alpha' < g(x) < \beta',$$

and this sub-set of E_1^0 is clearly $E_1^0 E_2^0$. $E_1^0 E_2^0$ is therefore not a null set. Further, E_1^0 and E_2^0 both being open sets, their common part $E_1^0 E_2^0$ is also open. Also, since an open set is the sum of a finite number or enumerably infinity of non-overlapping open intervals,

$$|E_1^0 E_2^0| > 0.$$

Again, since

$$E_1 E_2 \supset E_1^0 E_2^0,$$

we have

$$|E_1 E_2| \geqslant |E_1^0 E_2^0| > 0.$$

The author is indebted to the referee for some kind and valuable suggestions.

A POINT REPRESENTATION OF A SYSTEM OF SPACE CUBIC CURVES WHICH PASS THROUGH FOUR GIVEN POINTS AND WHOSE CHORDS BELONG TO A GIVEN TETRAHEDRAL COMPLEX

By V. C. MORTON and M. T. CHAPPLE (*Aberystwyth*)

[Received 30 April 1947]

1. Introduction

THE congruence of space cubic curves which pass through four given points and have a given line as chord has been discussed by Reye* and Sturm.† Stuyvaert‡ has given an algebraic representation. More recently the method of De Vries§ and Schaake‖ has been to set up in various ways a representation of cubics of the congruence by points of a plane.

In this paper we consider a wider system. In fact it seems to be simpler to look upon the congruence as a sub-family of a system of ∞^3 space cubics which pass through four given points and whose chords belong to a given tetrahedral complex of which the four points are the fundamental points. We show that this system admits a simple representation by the points of threefold space in such a way that the ∞^3 planes of space represent the ∞^3 congruences of the given chord type which belong to the system. Thus, as special cases, there arise plane representations of these congruences similar to those given in more recent papers. The representation also has the advantage that all degenerate cubics of the system are exhibited in an obvious manner. Once the representation is set up all the cubics of the system which obey a onefold or twofold condition are represented by the points of a representative surface or curve.

2. The point representation of cubics ρ of the system

Consider the reciprocal relations in threefold space S between points T and t defined by the equations

$$\text{(i)} \quad X_T = \frac{ax_t}{a+\theta}, \text{ etc.} \qquad \text{(ii)} \quad x_t = \frac{(a+\theta)}{a} X_T, \text{ etc.}$$

A point $t(x_t)$ of S may be regarded as the image of the unique space cubic curve ρ given by equations (i): this curve passes through $t(x_t)$

* (1) 206. † (2) 25–31. ‡ (3). § (4), (5), and (6). ‖ (7).

and the tetrad of reference 1234. On the other hand, equations (ii) define a complex of lines τ which have as images the points $T(X_T)$ of space S. The lines τ are obviously joins of two projectivities of points T and T' defined by the relations $X_T = aX'_T$, etc. Therefore the lines τ belong to a tetrahedral complex K for which the points 1234 are united points, and, further, chords of space cubic curves ρ belong to the complex K of lines.

Thus the ∞^3 points $t(x_t)$ of threefold space S can be regarded as the images of the complex of space cubic curves ρ which pass through the four united points of a tetrahedral complex of lines K and whose chords belong to the complex.

Further, $t(x_t)$ and ρ, and also $T(X_T)$ and τ, are in one-to-one correspondence in the sense that to a point $t(x_t)$ corresponds a unique cubic space curve ρ, while to the points $T(X_T)$ of this curve ρ correspond ∞^1 lines τ of K which are concurrent in the image point $t(x_t)$.*

3. The representation of cubic curves ρ which obey certain elementary conditions

If the image point $t(x_t)$ lies in a plane L, $(px) = 0$, then the corresponding cubic curve ρ meets L in t and two other points T_1 and T_2 which together form a triangle circumscribed to the complex conic in L and also self-conjugate with respect to the section by L of the quadric $(p/ax_t)X^2 = 0$. T_1, T_2 therefore lie on the line L, $(px) = 0$, M, $(px/a) = 0$, of K. Hence points of a plane L, $(px) = 0$, are images of the congruence of cubic curves ρ which have a common chord: the chord is the intersection of L, $(px) = 0$, and M, $(px/a) = 0$. Conversely, cubic curves ρ with a common chord are represented by points of a unique plane.

A cubic curve ρ which meets a chord α in two points T_1, T_2 is represented by the point of intersection of the tangents from T_1 and T_2 to the complex conic in the plane L corresponding to α.†

Hence the order of the representative curve of cubics ρ which obey any twofold condition is equal to the number of such cubic curves which have a given line of K as chord.

Image points t lying on a line τ of K represent those ∞^1 cubic

* The same relationships between points and lines, and between points and curves can be set up geometrically by considering the dual of a correspondence given by Sturm (8) 367.

† This is just the representation of the congruence of cubic curves with a given chord which is developed by De Vries in (5).

curves ρ which lie on the complex cone of the image point T on τ. Conversely, the image point of a cubic curve passing through T must lie on τ. Furthermore, a line of K which meets a cubic curve ρ must meet it twice. A line u of general position represents a pencil of cubic curves ρ lying on a quadric U containing u and the tetrad of points 1234.

Hence the order of the representative surface of cubic curves ρ which obey any onefold condition is equal to the number of such cubic curves which either pass through a given point T or lie on a quadric U.

Cubic curves ρ which touch a line α of the complex K are represented by the points of the complex conic of the corresponding image plane L, while those which meet a general line l are represented by points of a corresponding quadric λ which contains l and touches the faces of the tetrad 1234.

We have seen that lines τ of K can be regarded as the joins of two projectivities of points T and T', and that cubic curves ρ through a point T are represented by the line τ which corresponds to T; it follows that cubic curves ρ which touch a plane P are represented by the quartic developable surface π of the congruence of lines τ determined by points of P, while those which osculate P are represented by the edge of regression of π.

If the cubic curves ρ have also a fixed chord α, then the image points lie on the plane section of the developable by the plane L; the tricuspidal quartic curve obtained by Schaake, in (7) 780, is at once given. Moreover, the complex conic in L is then the fundamental conic on which the representation of Schaake is based.

The plane configuration cut out on a plane P by cubic curves ρ through the points 1, 2, 3, 4 and having a given chord α has been discussed in detail by Schaake and others. The contact curve is a cubic C^3 with a node at the point O where α meets P; the branch-point curve is the complex conic C^2 in the plane. C^3 touches C^2 at the three points K_1, K_2, K_3 where cubic curves of the congruence osculate P. A very simple method of dealing with the configuration is, however, noting that the three points where any cubic curve ρ meets P form a triangle circumscribed to C^2, to take O as (0, 1, 0) and C^2 as $(\theta^2, \theta, 1)$ in the plane. Then the parameters θ_1, θ_2, θ_3 of the points of contact of the sides of the triangle with C^2 may be shown to satisfy the involution relation $\theta_1\theta_2\theta_3 =$ constant. The main results emerge at once: thus the triple elements of the involution themselves

form a group of the involution, and it therefore follows that, if the tangents to C^2 at K_1, K_2, K_3 meet at P, Q, R, then a cubic curve ρ will pass through P, Q, R. In the representation given here the image of this cubic curve is the point of concurrence of the cuspidal tangents of the tricuspidal quartic curve referred to above. It may further be noted that the plane configuration is the dual of that cut out on a plane by the three chords, in the plane, of space cubic curves through three given points and having two given chords: for the three points of intersection of such cubics with the plane, in the latter congruence, move on a conic and form an involution g_3^2 upon it.

The order of the representative surface of those cubics ρ which have a proper tangent through a given point is four because four such cubics lie on quadric U. Cubic curves ρ passing through the unit point have equations $X = (a+\theta)/(a+\lambda)$, etc., whence, considering the intersections of a curve θ with a plane $H+\mu K = 0$ which passes through an arbitrary line $H = 0$, $K = 0$, we find that there are six cubic curves ρ lying on the complex cone of the unit point and having an osculating plane through the line $H = 0$, $K = 0$. Hence the order of the representative surface of those cubic curves ρ which have an osculating plane through a given line is six.*

4. There are two important types of general condition that, even in the case of a congruence of cubic curves ρ with a given chord, have received no attention. Firstly, the condition that a cubic curve ρ should meet once, or more often, a given curve C_p^n of order n and genus p, and, secondly, the condition that a cubic curve ρ should have various types of contact with a given surface† of order n.

5. Cubic curves ρ which meet a curve C_p^n

If T moves on a curve C_p^n, then the corresponding lines τ generate a scroll containing a pencil of directrix curves of order n and genus p. The scroll is therefore of order $2n$ and genus p.‡ Also each face of the tetrad 1234 is a tangent plane of multiplicity n of the scroll. The points of this scroll therefore represent the cubic curves ρ which meet the given curve C_p^n. The points of intersection of a generator τ of the scroll with the $2n-2$ other generators which meet τ are images of the

* Cf. (**7**).

† Cf. (**9**) 129, where the second type of condition is considered for the case of cubic curves through five given points.

‡ The numbers quoted in this paragraph are given in (**10**).

$2n-2$ cubic curves ρ each of which meets C_p^n at T and a second point of the curve. Cubic curves ρ which meet C_p^n twice are thus represented by the points of the double curve of the scroll, and this curve is of order $(n-1)(2n-1)-p$. Cubic curves ρ which touch C_p^n do so at its points of intersection with the torsal generators of the scroll. The number of such cubics is thus, in general, $4(n+p-1)$. Cubic curves ρ which meet C_p^n at three points are represented by the triple points of the scroll: in general their number is thus

$$\tfrac{2}{3}(n-1)(n-2)(2n-3)-2(n-2)p.$$

Finally it is possible to obtain sets of three collinear points on C_p^n such that a cubic curve ρ will pass through each pair of the three. These sets are given by those trisecants of C_p^n which belong to K. Their number is thus, in general,

$$\tfrac{2}{3}(n-1)(n-2)(n-3)-2(n-2)p.$$

6. Cubic curves ρ which have contact with a surface

If T moves on a general surface G of order n, then τ describes a congruence* Γ of order $3n$ and class n. Γ has, in general, no singular points; its only singular planes are the faces of the tetrad 1234; its only multiple lines are the edges of 1234, each of which is n-ple. We proceed to find the remaining characteristic numbers† of Γ.

The rank r of Γ is the number of pairs of rays of Γ which are concurrent on and coplanar with a given line l. Consider lines τ of K which meet an arbitrary line l. Points T which are images of τ then lie on a quadric surface U through l. Hence lines τ of Γ which meet l have image points which lie on a curve of order $2n$ and genus $(n-1)^2$ meeting l in n points. To each point of l corresponds one cubic curve ρ meeting the $2n$-ic in $3n$ points belonging to a g^1_{3n}. Similarly a pencil of planes through l determine on the $2n$-ic a g^1_n. r is then the number of pairs common to a g^1_{3n} and g^1_n. Hence $r = 2n(n-1)$ and $a = 2n(n+2)$. Since G has been supposed general, then $d = 6\,{}^nC_2$ and therefore $j = 2n(2n+1)$. It follows‡ that those cubic curves ρ which touch G are represented by the focal surface ϕ of Γ which is of order $2n(n+1)$ and class $2n(n-1)$. Cubic curves ρ which osculate G are represented by the points of the cuspidal curve f_3 on ϕ and the order μ_3 of this curve is the number of such cubic curves which have

* (8) 371.

† Vide (11) 74. We use the notation of this paper.

‡ (11) 79.

a given chord. Now the image point of the unique cubic curve ρ which has for chord the line L, $(px) = 0$, M, $(px/a) = 0$, and which passes through the point $T(x_1)$ is the point of intersection of L and the line τ corresponding to T. The cubic curve has therefore parametric equations

$$x = \frac{x_1(L_1 - aM_1)}{a+\theta}, \text{ etc.},$$

which, by change of parameter, may be written

$$x = \frac{x_1(L_1 - aM_1)}{\theta + L_1 - aM_1}.$$

Suppose now that this cubic curve touches, at the point (x_1), the surface G whose equation, in usual symbolic form, is

$$(px)^n = 0. \tag{1}$$

Then, on differentiating with respect to θ and putting $\theta = 0$ after substitution for (x), we find that the point of contact must satisfy (1) and also

$$(px)^{n-1}[px/L - aM] = 0,$$

which, by using (1), may be written

$$(px)^{n-1}[px\{(b+c+d)L^2 - (cd+db+bc)LM + bcdM^2\}] = 0. \tag{2}$$

Thus, cubic curves ρ with a given chord touch G at points of the curve of intersection of (1) and (2), a curve of order $n(n+2)$. If the cubic curve osculates G, we find, by further differentiation, that the points of osculation must lie on (1), (2) and also on the surface

$$\begin{aligned}(n-1)(px)^{n-2}[px\{(b+c+d)L^2 - (cd+dc+bc)LM + bcdM^2\}]^2 - \\ - 2(px)^{n-1}[px\{(b+c+d)L^2 - (cd+db+bc)LM + bcdM^2\}^2] = 0,\end{aligned} \tag{3}$$

a surface of order $n+4$.

But the given chord $L = M = 0$ is double on (2) and quadruple on (3), so that the residual curve of intersection of (2) and (3) is of order n^2+6n. Furthermore, the four lines $L - aM = 0$, $x = 0$, etc., meet the surfaces (2) and (3) at the $4n$ points where these lines meet (1). Thus the number of points of intersection of (1), (2), (3) which lead to points of osculation is $n(n^2+6n-4)$. This is then the number of cubic curves ρ which have a given chord and osculate G and so is also the order μ_3 of the image curve f_3 of cubic curves ρ which osculate G.

Since f_3 is the cuspidal curve of ϕ, then* the rank c of ϕ is $n(n^2+2)$.

* (**10**) 80.

The congruence Γ is now completely determined by its five known characteristics m, n, a, j, c.

By use of the results given in (**10**) 80–2, we can now, after some reduction, give the following numbers:

(i) cubic curves ρ which are bitangent to G are represented by the points of the double curve of ϕ: this curve is of order

$$2n(n-1)(n^2+2n-2);$$

(ii) those cubic curves ρ which have four-point contact with G are represented by the stationary points of the cuspidal curve of ϕ: their number μ_4 is $4n(n-1)(3n-2)$;

(iii) those cubic curves ρ which touch G at one point and osculate it at another are represented by the stationary points on the double curve of ϕ: their number μ_{31} is $2n(n-1)(n^3+8n^2-20n+12)$;

(iv) those cubic curves ρ which are tritangent to G are represented by the triple points on the double curve of ϕ: their number is

$$\tfrac{1}{3}n(n-1)(n-2)(n^3+3n^2-8n+6).$$

7. Finally, by finding the number of points of intersection of known loci, we can find the number of cubic curves ρ which obey combinations of such conditions as we have considered.

REFERENCES

1. Reye, *Geometrie der Lage*, 3, 4. Aufl. (Leipzig, 1910).
2. Sturm, *Die Lehre von den geometrischen Verwandtschaften*, iv (Leipzig and Berlin, 1909).
3. Stuyvaert, 'Congruence de cubiques gauches', *Acad. Roy. de Belgique* (2), 4 (1920), 65–80.
4. De Vries, 'A congruence (1.0) of twisted cubics', *K. Akad. Amsterdam, Proc.* 26 (1923), 126.
5. —— 'Linear congruences of twisted cubics that cut at least one fixed line twice', ibid. 31 (1928), 629.
6. —— 'Eine Abbildung der Kongruenz der kubischen Raumkurven durch vier Punkte, welche eine vorgegebene Gerade zweimal treffen', ibid. 37 (1934), 129.
7. Schaake, 'A representation of the congruence of twisted cubics which pass through four given points and which cut a given straight line twice', ibid. 29 (1926), 776.
8. Sturm, *Liniengeometrie*, i. (Leipzig, 1892).
9. J. R. Conner, 'Basic systems of rational normal curves', *American J. of Math.* 32 (1910), 115.
10. Edge, *Ruled Surfaces* (Cambridge, 1931).
11. L. Roth, 'Line congruences in three dimensions', *Proc. London Math. Soc.* (2), 32 (1931), 72.

A CLASS OF COMPLEX MARKOFF CHAINS

By P. A. P. MORAN (*Oxford*)

[Received 24 June 1947]

1. SUPPOSE that x is a random variable which can take any one of a finite set of values and suppose that in repeated trials the probability that it takes any particular value at a given trial depends on the previous results. This situation is described as a *Markoff chain* and, if the probabilities at any given trial depend only on the results of the immediately preceding trial, the chain is said to be *simple*. If they depend also on earlier results, the chain is said to be *complex*. Simple chains have an extensive literature,* but complex chains have been very little studied. In this paper I discuss a class of complex Markoff chains which were suggested by the study of empirical time series. The equivalent of the 'central limit' theorem is proved for the sums $\sum_1^n x_i$ and $T = \sum_1^n x_i y_i$, where $\{x_i\}$ and $\{y_i\}$ are two statistically independent chains of the type considered. The asymptotic distribution of the statistic T is discussed in relation to a rough test for the existence of a correlation between two such series.

2. Suppose that x_t $(t = 1, 2, \ldots)$ are random variables which can take the values ± 1, and let $x_1, \ldots, x_k$ be fixed. Then, for any given sequence

$$x_{t+1}, \ldots, x_{t+k} \quad (t \geqslant 0),$$

we suppose that the probability that x_{t+k+1} is equal to 1 is

$$P_1^{t+k+1} = \tfrac{1}{2} - a_1 x_{t+k} - \ldots - a_k x_{t+1}$$

and that the probability that it is equal to -1 is

$$P_{-1}^{t+k+1} = 1 - P_1^{t+k+1} = \tfrac{1}{2} + a_1 x_{t+k} + \ldots + a_k x_{t+1},$$

where the a_i are real constants such that

$$|a_1| + \ldots + |a_k| = \tfrac{1}{2} - \delta < \tfrac{1}{2}. \tag{1}$$

P_1^m and P_{-1}^m are therefore defined as the probabilities that $x_m = 1$ and -1 respectively when $x_{m-1}, \ldots, x_{m-k}$ are known. The relation above then defines a probability process which generates the series from x_{k+1} onwards, assuming that $x_1, \ldots, x_k$ are known and fixed. Now

* Hostinsky (**1**), Fréchet (**2**).

define p_1^m and p_{-1}^m ($m > k$) as the probabilities that $x_m = 1$ and -1 when we are given the values of $x_1,\ldots,x_k$ only. For $m > k$, define

$$2Q_m = 2p_1^m - 1 = 1 - 2p_{-1}^m = p_1^m - p_{-1}^m$$

as the expectation of x_m, given $x_1,\ldots,x_k$, and write

$$2Q_m = -2E(a_1 x_{m-1} + \ldots + a_k x_{m-k}), \tag{2}$$

where the expectation of the expression in brackets assumes a knowledge of $x_1,\ldots,x_k$ only. For $1 \leqslant m \leqslant k$ we write $Q_m = \pm\frac{1}{2}$ according as $x_m = \pm 1$.

From (2) we see that

$$Q_m + 2a_1 Q_{m-1} + \ldots + 2a_k Q_{m-k} = 0 \tag{3}$$

for $m > k$. Moreover, Q_m is given for $m = 1,\ldots,k$, and the later values are determined uniquely in terms of these by (3).

From (1), since $\delta > 0$, the characteristic equation

$$t^k + 2a_1 t^{k-1} + \ldots + 2a_k = 0 \tag{4}$$

will have all its roots inside the circle $|t| = 1$, and, in fact, in or on the circle $|t|^k = 1 - 2\delta$. Let these roots be $w_1,\ldots,w_p$ with multiplicities $\pi_1,\ldots,\pi_p$. Then the solution of (3) can be written in the form

$$Q_m = \sum_{s=1}^{p} H_s^{\pi_s - 1}(m) w_s^m, \tag{5}$$

where $H_s^r(m)$ is a polynomial in m of degree r whose coefficients are real for real roots, and for complex roots, which occur in conjugate pairs, are such that the sum of the corresponding two terms is real. The number of arbitrary constants is therefore $\pi_1 + \ldots + \pi_p = k$, and they can be chosen to make the Q_m defined by equation (5) satisfy the conditions imposed by fixing $x_1,\ldots,x_k$.

It follows from the condition we have imposed on the roots of equation (4) that $Q_m \to 0$ as m increases, and so p_1^m and p_{-1}^m both tend to $\frac{1}{2}$. The implication of this is that two x's sufficiently far apart tend to behave as if they were effectively independent. By a more elaborate method we can see that the joint probability distribution of x_m and x_{m+s}, where s is fixed, will tend to a definite probability distribution independent of $x_1,\ldots,x_k$ when m increases.

Similarly, it is clear that it is possible to show that the joint probability distribution of

$$(x_m, x_{m+s_1},\ldots,x_{m+s_t}) \quad (k < m < m+s_1 < \ldots < m+s_t)$$

tends to a definite distribution independent of m when m increases.

We may therefore suppose that we have a sequence of random variables $(\ldots x_{-1}, x_0, x_1, \ldots)$ forming such a Markoff chain and started up so long ago that the joint distribution of $(x_m, x_{m+s_1}, \ldots, x_{m+s_l})$, can be regarded as independent of m, and therefore the serial covariance

$$c_s = E(x_m x_{m+s}) = E(x_m x_{m-s}) = c_{-s}$$

can also be taken as a function of s only.

In what follows we therefore drop the condition that $x_1, \ldots, x_k$ are fixed and instead suppose the series to have been started up a long time previously. Since

$$E(x_m^2) = E(x_{m+t}^2) = 1,$$

the serial covariance c_s is also the serial correlation coefficient, which we write r_s.

3. Now consider the expectation of

$$-2(a_1 x_k + \ldots + a_k x_1) x_{k-p}, \tag{6}$$

where $p \geqslant 0$. Multiplying out and taking the expectation, we see that this is equal to

$$-2(a_1 r_p + \ldots + a_k r_{p-k+1}).$$

On the other hand, (6) is also equal to

$$x_{k-p}\{\mathrm{pr}(x_{k+1} = 1) - \mathrm{pr}(x_{k+1} = -1)\}, \tag{7}$$

where the probabilities are calculated on the assumption that $x_1, \ldots, x_k$ are known. If we take the average of (7) over all possible sets of values with their correct relative frequencies, we see that this is equal to r_{p+1}, and so we have

$$r_{p+1} + 2a_1 r_p + \ldots + 2a_k r_{p-k+1} = 0 \tag{8}$$

for $p \geqslant 0$. This equation is not true for $p < 0$. Thus r_m satisfies the same difference equation as Q_m, and tends to zero, as m tends to infinity, at least as fast as $(1-2\delta)^{m/k}$, for the roots of (4) lie in or on the circle $|t|^k = 1-2\delta$. If we put $p = 0, 1, \ldots, k-1$ in (8), we obtain a set of k equations for $r_1, \ldots, r_k$, and we shall show that these equations are sufficient to determine $r_1, \ldots, r_k$ uniquely. We can then form $r_{k+1}, \ldots$ by the repeated use of (8). The k equations are

$$\left.\begin{aligned} r_1 + 2a_1 r_0 + 2a_2 r_1 + \ldots + 2a_k r_{k-1} &= 0 \\ r_2 + 2a_1 r_1 + 2a_2 r_0 + \ldots + 2a_k r_{k-2} &= 0 \\ \cdot \quad \cdot \quad \cdot \quad

Writing $\mathbf{r}$ for the column matrix with elements $(r_1,\ldots,r_k)$ and $\mathbf{a}$ for the column matrix with elements $(2a_1,\ldots,2a_k)$, we can write these equations (since $r_0 = 1$) as

$$\mathbf{Ar} = -\mathbf{a},$$

where $\mathbf{A} = \mathbf{1}+\mathbf{B}+\mathbf{C}$, and $\mathbf{1}$ is a unit $k\times k$ matrix, and $\mathbf{B}$ and $\mathbf{C}$ are the matrices

$$\begin{bmatrix} 2a_2 & 2a_3 & \ldots & 2a_{k-1} & 2a_k & 0 \\ 2a_3 & 2a_4 & \ldots & 2a_k & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 2a_k & 0 & \ldots & 0 & 0 & 0 \\ 0 & 0 & \ldots & 0 & 0 & 0 \end{bmatrix}$$

and

$$\begin{bmatrix} 0 & 0 & \ldots & 0 & 0 & 0 \\ 2a_1 & 0 & \ldots & 0 & 0 & 0 \\ 2a_2 & 2a_1 & \ldots & 0 & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 2a_{k-1} & 2a_{k-2} & \ldots & 2a_2 & 2a_1 & 0 \end{bmatrix}.$$

We must show that $|\mathbf{A}|$ is not zero. If it were zero, there would exist numbers $\lambda_1,\ldots,\lambda_k$ (not all zero) such that, if a_{ij} is a typical element of the matrix $\mathbf{A}$, then

$$\sum_{j=1}^{k} \lambda_j a_{ij} = 0.$$

The pth row of the matrix $\mathbf{A}$ contains each a_s just once, except for a_p which is absent.

Let λ_p be a λ whose modulus is not less than any of the others. Then

$$\lambda_p + 2\sum_{i=1}^{k}{}' \lambda_{m(i)} a_i = 0,$$

where $\lambda_{m(i)}$ is the λ corresponding to the column in which a_i appears and the dash on the sign of summation indicates that we omit the term corresponding to a_p. But $\lambda_p \neq 0$, and so we have

$$1 + 2\sum_{i=1}^{k}{}' \frac{\lambda_{m(i)} a_i}{\lambda_p} = 0,$$

which is impossible since $|\lambda_{m(i)}| \leqslant |\lambda_p|$ and $2\sum_{1}^{k}{}' |a_i| \leqslant 1-2\delta$. The equations (9) are therefore sufficient to determine $r_1,\ldots,r_k$. I shall later discuss another method of finding the r's.

4. Now write $X_m = x_m + 2a_1 x_{m-1} + \ldots + 2a_k x_{m-k}$. Then the sequence of X's is a sequence of random variables defined in terms of the x's. Also $|X_m| \geqslant 2\delta$, and so

$$d = E(X_m^2) \geqslant 4\delta^2 > 0; \tag{10}$$

if desired, d can be evaluated in terms of the r's.

Also, for $s > 0$,

$$E(X_{m+s} X_m) = E\{(x_{m+s} + 2a_1 x_{m+s-1} + \ldots + 2a_k x_{m+s-k}) \times \\ \times (x_m + 2a_1 x_{m-1} + \ldots + 2a_k x_{m-k})\},$$

and, expanding the product according to the second of these factors and taking expectations, we get

$$E(X_{m+s} X_m) = 0.$$

Nevertheless, the sequence of X's is not a sequence of independent random variables, for the distribution of X_m depends on the values of $x_{m-1}, \ldots, x_{m-k}$. The X's bear an analogy to the independent random variables used in stochastic difference equations but they are not independent of the previous x's, and it is therefore not useful to expand the x's in terms of them. Thus, although the present class of Markoff chains shows a very strong analogy in behaviour to the solutions of stochastic difference equations, the proofs employed must be somewhat different.

5. I now prove that the 'central limit' theorem holds for the sum $S = \sum_1^n x_i$. In the first place

$$E(S) = \sum_1^n E(x_i) = 0.$$

Consider

$$E(S^2) = E(x_1 + \ldots + x_n)^2 = n + 2\sum_{i=1}^{n-1} (n-i) r_i.$$

Then

$$n^{-1} E(S^2) \to 1 + 2\sum_{i=1}^{\infty} r_i. \tag{11}$$

Here we are using the consistence theorem for Cesàro summation. The right-hand side is finite because the series is majorized by a convergent geometric series. We must show that the right-hand side of (11) is not zero. Consider the sum

$$S_1 = (1 + 2a_1 + \ldots + 2a_k) \sum_1^n x_i$$

$$= \sum_{i=k+1}^n X_i + P \quad \text{(say)},$$

where P is the sum of $2k$ terms in $x_1, x_2, \ldots, x_k, x_{n-k+1}, \ldots, x_n$. Then $E\Big(\sum_{i=k+1}^{n} X_i\Big)^2 = (n-k)d$ and, when n is large, this increases asymptotically as nd.

Now $E(P^2)$ is constant and $E\Big(P\sum_{k+1}^{n} X_i\Big)$ can be shown to be bounded independently of n because the sequence $\{r_i\}$ is majorized by a convergent geometric series. $E(S_1^2)$ therefore increases as n when n is large, and

$$n^{-1}(1+2a_1+\ldots+2a_k)^2 E\Big(\sum_1^n x_i\Big)^2 \to d \neq 0$$

when n increases, and the constant d can be evaluated by expanding (10). I now use a theorem due to S. N. Bernstein,* which, on simplification, may be stated as follows:

THEOREM. *If $\Sigma_n = x_1+\ldots+x_n$ is a sum of dependent variables such that*

(1) $E(x_i) = 0$ *for all* i,

(2) $B_n = E(\Sigma_n^2) > Mn$ *for all* n, *where* M *is a constant greater than zero*,

(3) $|x_i|$ *is bounded by a constant independent of* i,

(4) *there exist positive numbers* ϵ *and* ρ $(\rho < \frac{1}{3})$ *such that the range of variation of* $E(x_i \mid x_k)$ *for all possible values of* x_k *is less than* $e^{-n\epsilon}$ *whenever* $|i-k| > n^\rho$;

then

$$\mathrm{pr}\{t_0 B_n^{\frac{1}{2}} \leqslant \Sigma_n < t_1 B_n^{\frac{1}{2}}\} \to \frac{1}{\sqrt{(2\pi)}} \int_{t_0}^{t_1} e^{-\frac{1}{2}t^2}\, dt \text{ uniformly in } t_0 \text{ and } t_1.$$

Apply this theorem to the complex Markoff chain defined above. By multiplying the x's by a constant it can be seen that in condition (4) it is sufficient to ensure that the variation of $E(x_i \mid x_k)$ is less than $Ke^{-n\epsilon}$, where K is some positive constant. We conclude from what we have seen before that $E(S^2) > Mn$, where M is a positive constant and $|x_i| = 1$, so that condition (3) is satisfied. Moreover,

$$r_{i-n} = E\{x_n E(x_i \mid x_n)\} = \tfrac{1}{2}\{E(x_i \mid x_n = 1) - E(x_i \mid x_n = -1)\},$$

and so the range of variation of $E(x_i \mid x_n)$ is $2r_{i-n}$. But $|2r_{i-n}|$ is less than $K_1 \rho_1^{|i-n|}$, where K_1 is a positive constant and $0 < \rho_1 < 1$, and so we can choose a constant $\epsilon > 0$ such that $|2r_{i-n}| < K_2 e^{-n\epsilon}$, where $K_2 > 0$.

* (4) 36.

Thus the conditions of the theorem are satisfied and

$$\text{pr}\{t_0(E(S^2))^{\frac{1}{2}} \leqslant S < t_1(E(S^2))^{\frac{1}{2}}\} \to (2\pi)^{-\frac{1}{2}} \int_{t_0}^{t_1} e^{-\frac{1}{2}t^2}\, dt$$

uniformly in t_0 and t_1.

6. We next consider two such series, the second being generated by the relation

$$\text{pr}\{y_{l+1} = -1\} = 1 - \text{pr}\{y_{l+1} = 1\} = \tfrac{1}{2} + A_1 y_l + \ldots + A_l y_1,$$

where $\sum |A_i| = \frac{1}{2} - \delta' < \frac{1}{2}$. Suppose this series to have been started up a long time before and write

$$R_s = E(y_m y_{m+s}) = E(y_m y_{m-s}).$$

I construct a test for the independence of two observed series by considering the distribution of

$$T = \sum_{i=1}^{n} x_i y_i.$$

Then, if there is complete independence,

$$E(T) = 0,$$

and

$$\begin{aligned} E(T^2) &= E\Big(\sum_{i=1}^{n} x_i y_i\Big)^2 \\ &= n + 2E\Big(\sum_{t=1}^{n-1} \sum_{s=1}^{n-t} x_t x_{t+s} y_t y_{t+s}\Big) \\ &= n + 2\sum_{s=1}^{n-1} (n-s) r_s R_s, \end{aligned}$$

and

$$n^{-1}E(T^2) \to 1 + 2\sum_{s=1}^{\infty} r_s R_s = R \quad \text{(say)}. \tag{11a}$$

The series on the right converges, being dominated by a convergent geometric series, and we must prove that its sum is not equal to zero.

I have already said that the behaviour of these chains mimics the behaviour of the solutions of stochastic difference equations of the form

$$Z(t) + 2a_1 Z(t-1) + \ldots + 2a_k Z(t-k) = \eta(t),$$

where t takes integral values and the $\eta(t)$ $(t = 0, \pm 1, \ldots)$ are independent random variables with zero mean and finite second moments. If this equation generates a stationary stochastic process, it is known that $Z(t)$ can be expanded in a convergent infinite series in the $\eta(t)$ of the form

$$Z(t) = \sum_{i=0}^{\infty} \alpha_i \eta(t-i). \tag{12}$$

This is not possible in the present case for now corresponding to the $\eta(t)$ are the quantities

$$X_m = x_m + 2a_1 x_{m-1} + \ldots + 2a_k x_{m-k}$$

and this has a distribution depending on $x_{m-1}, \ldots, x_{m-k}$. In spite of this, the coefficients α_i in the expansion (12) have analogues in the present case which play an essential part in the theory. We define the sequence $\alpha_0, \alpha_1, \ldots$ by the relations

$$\left.\begin{aligned} 1 &= \alpha_0 \\ 0 &= \alpha_1 + 2a_1\alpha_0 \\ 0 &= \alpha_2 + 2a_1\alpha_1 + 2a_2\alpha_0 \\ &\cdots\cdots\cdots\cdots \\ 0 &= \alpha_{k-1} + 2a_1\alpha_{k-2} + \ldots + 2a_{k-1}\alpha_0 \\ 0 &= \alpha_{k+s} + 2a_1\alpha_{k+s-1} + \ldots + 2a_k\alpha_s \quad (s \geqslant 0) \end{aligned}\right\}. \tag{13}$$

Clearly $|\alpha_i| \leqslant 1$ for $0 \leqslant i \leqslant k$, and, for $i > k$, $|\alpha_i| < b^{i-k}$, say, where $0 < b < 1$. The α_i are therefore majorized by a convergent geometric series.

I now prove that

$$r_s = \Big(\sum_{i=0}^{\infty} \alpha_i^2\Big)^{-1} \Big(\sum_{i=0}^{\infty} \alpha_i \alpha_{i+s}\Big). \tag{14}$$

This is done by showing that the right-hand side of (14) satisfies the equations (8). Since the α_i satisfy (13), this is certainly true for $s \geqslant k$. Now suppose that $0 \leqslant p < k-1$. Then we want to show that

$$r_{p+1} + 2a_1 r_p + \ldots + 2a_{p+1} r_0 + 2a_{p+2} r_1 + \ldots + 2a_k r_{k-p-1} = 0,$$

is satisfied by (14). Inserting (14) in this and omitting the factor $\Big(\sum_{i=0}^{\infty} \alpha_i^2\Big)^{-1}$ we must show that

$$-\sum_{i=0}^{\infty} [\alpha_i \alpha_{i+p+1} + \ldots + 2a_{p+1}\alpha_i^2 + \ldots + 2a_k \alpha_i \alpha_{i+k-p-1}] = 0,$$

and, after some rearrangement, this becomes

$$\begin{aligned} &\alpha_0(\alpha_{p+1} + 2a_1\alpha_p + \ldots + 2a_{p+1}\alpha_0) + \\ &+\alpha_1(\alpha_{p+2} + 2a_1\alpha_{p+1} + \ldots + 2a_{p+2}\alpha_0) + \\ &\cdots\cdots\cdots\cdots \\ &+\alpha_{k-p-1}(\alpha_k + 2a_1\alpha_{k-1} + \ldots + 2a_k\alpha_0) + \\ &+\alpha_{k-p}(\alpha_{k+1} + 2a_1\alpha_k + \ldots + 2a_k\alpha_1) + \\ &\cdots\cdots\cdots\cdots \end{aligned}$$

which vanishes because of (13).

Similarly, we can show that

$$R_s = \Big(\sum_{i=0}^{\infty} \beta_i \beta_{i+s}\Big)\Big(\sum_{i=0}^{\infty} \beta_i^2\Big)^{-1},$$

where the β_i are the solutions of the corresponding set of equations for the chain in $\{y_i\}$. We can now prove that

$$R = 1+2\sum_{s=1}^{\infty} r_s R_s \neq 0.$$

For

$$R = \Big(\sum_{i=0}^{\infty} \alpha_i^2 \sum_{j=0}^{\infty} \beta_j^2\Big)^{-1}\Big(\sum_{i=0}^{\infty} \alpha_i^2 \sum_{j=0}^{\infty} \beta_j^2 + 2\sum_{s=1}^{\infty}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \alpha_i \alpha_{i+s} \beta_j \beta_{j+s}\Big)$$

$$= \Big(\sum_{i=0}^{\infty} \alpha_i^2 \sum_{j=0}^{\infty} \beta_j^2\Big)^{-1}\{\alpha_0^2\beta_0^2 + (\alpha_0\beta_1+\alpha_1\beta_0)^2 + (\alpha_0\beta_2+\alpha_1\beta_1+\alpha_2\beta_0)^2 + \ldots\}$$

$$\neq 0.$$

By using the α_i we can similarly give another proof that $1+2\sum_{i=1}^{\infty} r_i$ is not zero, using the fact that $\sum \alpha_i \neq 0$, which follows from (13).

Write $$u_i = x_i y_i.$$

Then $$E(T) = E\Big(\sum_1^n u_i\Big) = 0,$$

and $n^{-1}E(T^2)$ tends to a non-zero constant. Also $|u_i| = 1$ for any integer q.

That $\epsilon > 0$ can be chosen so that $E(u_i \mid u_k)$ is always less than $Ke^{-|i-k|\epsilon}$, where K is a positive constant, can be shown by an argument similar to that for $E(x_i \mid x_u)$. Then

$$\mathrm{pr}\{t_0(E(T^2))^{\frac{1}{2}} \leqslant T < t_1(E(T^2))^{\frac{1}{2}}\} \to (2\pi)^{-\frac{1}{2}} \int_{t_0}^{t_1} e^{-\frac{1}{2}t^2}\, dt$$

uniformly in t_0 and t_1.

Similar theorems applying to the solutions of stochastic difference equations were discussed in another paper (**3**). The present investigation was undertaken with the idea that by calculating the mean of a time series and considering only whether the individual values were greater or less than this, a quicker method could be devised for determining whether two empirical series were correlated or not. This does not seem to be true, for in order to calculate $E(T^2)$ it is necessary to estimate $a_1, \ldots, a_k$ and $A_1, \ldots, A_l$ and then calculate the corresponding value of $E(T^2)$. It is not sufficient to calculate the

empirical serial correlations and insert them in (11 a), because their sampling errors are so large that we cannot expect the resulting series to converge.

REFERENCES

1. B. Hostinsky, *Méthodes générales du Calcul des Probabilités* (Paris, 1931).
2. M. Fréchet, *Recherches théoriques modernes sur le Calcul des Probabilités*, Vol. 2 (Paris).
3. P. A. P. Moran, 'Some theorems on time series': *Biometrika* **34** (1947), 281.
4. S. N. Bernstein, *Math. Ann.* **97** (1927), 1.

THE HAMILTONIAN DERIVATIVES OF A CLASS OF FUNDAMENTAL INVARIANTS

By H. A. BUCHDAHL (*Hobart*)

[Received 27 August 1947]

1. The Hamiltonian derivatives of fundamental invariants play a prominent part in the Theory of Relativity* since they are symmetrical fundamental tensors the divergence of which vanishes identically. In the present paper I consider the class of fundamental invariants which contain the derivatives of the $g_{\mu\nu}$ only up to the second order. In particular I shall establish amongst other results the following *explicit* formula for the Hamiltonian derivative $hK/hg_{\mu\nu}$ of such an invariant K

$$\frac{hK}{hg_{\mu\nu}} = Z^{\mu\nu\sigma\rho}{}_{;\sigma\rho} - \tfrac{2}{3}B^{\mu}{}_{.\alpha\sigma\rho}Z^{\nu\alpha\sigma\rho} + \tfrac{1}{2}g^{\mu\nu}K, \tag{1.1}$$

where $Z^{\mu\nu\sigma\rho}$ is the tensor $\partial K/\partial g_{\mu\nu,\sigma\rho}$ symmetrical in μ, ν and in σ, ρ, which may generally be determined by inspection when the form of K is known explicitly. Subscripts after a comma or a semicolon denote ordinary and covariant differentiation respectively, whilst the order of the subscripts of the covariant curvature tensor $B_{\mu\nu\sigma\rho}$ is indicated in (4.3) below. We can then verify directly that the divergence of the right-hand side of (1.1) vanishes identically; and the derivation of the Hamiltonian derivatives of various specific invariants is a matter of ease.

2. Since, by hypothesis, K contains the derivatives of the $g_{\mu\nu}$ only up to the second order, it follows from a well-known result of the calculus of variations that

$$P^{\mu\nu} \equiv \frac{hK}{hg_{\mu\nu}} = \frac{1}{\gamma}\left\{\frac{\partial(\gamma K)}{\partial g_{\mu\nu}} - \frac{\partial}{\partial x_\alpha}\frac{\partial(\gamma K)}{\partial g_{\mu\nu,\alpha}} + \frac{\partial^2}{\partial x_\alpha \partial x_\beta}\frac{\partial(\gamma K)}{\partial g_{\mu\nu,\alpha\beta}}\right\}, \tag{2.1}$$

where γ is written for $\sqrt{(-g)}$.

Now, if $\theta(K)$ is an arbitrary function of K (satisfying such conditions of continuity and differentiability as are implied in the following argument), then $\theta(K)$ is itself a fundamental invariant; and its Hamiltonian derivative is given by (2.1) if K be replaced by $\theta(K)$.

Write

$$\theta' = \frac{d\theta(K)}{dK}, \qquad \theta'' = \frac{d^2\theta(K)}{dK^2},$$

* Eddington (**1**), § 61, ff.

and so on. Then from (2.1) we have

$$\frac{h\{\theta(K)\}}{hg_{\mu\nu}} = \tfrac{1}{2}g^{\mu\nu}\theta+\theta'\left[\frac{\partial K}{\partial g_{\mu\nu}}-\left(\frac{\partial K}{\partial g_{\mu\nu,\alpha}}\right)_{,\alpha}-\frac{\gamma_{,\alpha}}{\gamma}\frac{\partial K}{\partial g_{\mu\nu,\alpha}}+\right.$$
$$\left.+\left(\frac{\partial K}{\partial g_{\mu\nu,\alpha\beta}}\right)_{,\alpha\beta}+\frac{2\gamma_{,\alpha}}{\gamma}\left(\frac{\partial K}{\partial g_{\mu\nu,\alpha\beta}}\right)_{,\beta}+\frac{\gamma_{,\alpha\beta}}{\gamma}\frac{\partial K}{\partial g_{\mu\nu,\alpha\beta}}\right]+$$
$$+\theta''\left[-K_{,\alpha}\frac{\partial K}{\partial g_{\mu\nu,\alpha}}+K_{,\alpha\beta}\frac{\partial K}{\partial g_{\mu\nu,\alpha\beta}}+2K_{,\alpha}\left(\frac{\partial K}{\partial g_{\mu\nu,\alpha\beta}}\right)_{,\beta}+\frac{2\gamma_{,\beta}}{\gamma}K_{,\alpha}\frac{\partial K}{\partial g_{\mu\nu,\alpha\beta}}\right]+$$
$$+\theta'''K_{,\alpha}K_{,\beta}\frac{\partial K}{\partial g_{\mu\nu,\alpha\beta}}. \quad (2.2)$$

$P^{\mu\nu}$ is then given by (2.2) if we choose $\theta(K) = K$; i.e.

$$P^{\mu\nu} = \tfrac{1}{2}g^{\mu\nu}K+\frac{\partial K}{\partial g_{\mu\nu}}-\left(\frac{\partial K}{\partial g_{\mu\nu,\alpha}}\right)_{,\alpha}-\frac{\gamma_{,\alpha}}{\gamma}\frac{\partial K}{\partial g_{\mu\nu,\alpha}}+$$
$$+\left(\frac{\partial K}{\partial g_{\mu\nu,\alpha\beta}}\right)_{,\alpha\beta}+\frac{2\gamma_{,\alpha}}{\gamma}\left(\frac{\partial K}{\partial g_{\mu\nu,\alpha\beta}}\right)_{,\beta}+\frac{\gamma_{,\alpha\beta}}{\gamma}\frac{\partial K}{\partial g_{\mu\nu,\alpha\beta}}. \quad (2.3)$$

Now the right-hand side of (2.2) is a tensor. Since $\theta(K)$ is an arbitrary function of its argument,

$$K_{,\alpha}K_{,\beta}\frac{\partial K}{\partial g_{\mu\nu,\alpha\beta}} = K_{;\alpha}K_{;\beta}\frac{\partial K}{\partial g_{\mu\nu,\alpha\beta}}$$

must be a tensor. Hence $\dfrac{\partial K}{\partial g_{\mu\nu,\alpha\beta}}$, which is symmetrical in α, β (and in μ, ν) is a tensor. We write

$$\frac{\partial K}{\partial g_{\mu\nu,\alpha\beta}} = Z^{\mu\nu\alpha\beta}, \quad (2.4)$$

where
$$Z^{\mu\nu\alpha\beta} = Z^{\mu\nu\beta\alpha} = Z^{\nu\mu\alpha\beta}. \quad (2.41)$$

In the same way the coefficient of θ'' in (2.3), i.e.

$$-K_{\alpha}\frac{\partial K}{\partial g_{\mu\nu,\alpha}}+K_{,\alpha\beta}Z^{\mu\nu\alpha\beta}+2K_{,\alpha}Z^{\mu\nu\alpha\beta}{}_{,\beta}+\frac{2\gamma_{,\beta}}{\gamma}K_{,\alpha}Z^{\mu\nu\alpha\beta} \quad (2.5)$$

is a tensor. But

$$K_{,\alpha\beta} = (K_{;\alpha})_{,\beta} = K_{;\alpha\beta}+\{\alpha\beta,\sigma\}K_{;\sigma}, \quad (2.6)$$
$$\gamma_{,\beta}/\gamma = \{\beta\sigma,\sigma\}, \quad (2.61)$$

and

$$Z^{\mu\nu\alpha\beta}{}_{,\beta}$$
$$= Z^{\mu\nu\alpha\beta}{}_{;\beta}-\{\sigma\beta,\mu\}Z^{\sigma\nu\alpha\beta}-\{\sigma\beta,\nu\}Z^{\mu\sigma\alpha\beta}-\{\sigma\beta,\alpha\}Z^{\mu\nu\sigma\beta}-\{\sigma\beta,\beta\}Z^{\mu\nu\alpha\sigma}. \quad (2.62)$$

Substituting in (2.5), using (2.41), and changing dummy indices we get

$$K_{;\alpha\beta}Z^{\mu\nu\alpha\beta}+2K_{;\alpha}Z^{\mu\nu\alpha\beta}{}_{;\beta}-$$
$$-K_{;\alpha}\left(\frac{\partial K}{\partial g_{\mu\nu,\alpha}}+2\{\sigma\rho,\mu\}Z^{\nu\sigma\rho\alpha}+2\{\sigma\rho,\nu\}Z^{\mu\sigma\rho\alpha}+\{\sigma\rho,\alpha\}Z^{\mu\nu\sigma\rho}\right). \quad (2.7)$$

Hence the expression

$$Y^{\mu\nu\alpha}=\frac{\partial K}{\partial g_{\mu\nu,\alpha}}+2\{\sigma\rho,\mu\}Z^{\nu\sigma\rho\alpha}+2\{\sigma\rho,\nu\}Z^{\mu\sigma\rho\alpha}+\{\sigma\rho,\alpha\}Z^{\mu\nu\sigma\rho} \quad (2.8)$$

is a tensor.

3. Now without loss of generality we may for the time being (§§ 3–7) restrict ourselves to the case in which K has the form

$$K=\Gamma^{\mu_1\nu_1\ldots\rho_m}\prod_{s=1}^{m}B_{\mu_s\nu_s\sigma_s\rho_s}, \quad (3.1)$$

where $\Gamma^{\mu_1\nu_1\ldots\rho_m}$ is a tensor of rank $4m$ made up entirely of the product of $2m$ components of the metrical tensor. For the most general second-order invariant is of the form

$$K=F(K_1,K_2,\ldots), \quad (3.11)$$

where F is some function of the K_i, each of the latter being of the form (3.1). When the Hamiltonian derivatives of these are known, that of K follows at once, and is given by (1.1) (as is seen in § 8).

Fundamental *relative* invariants may also be dealt with by the present method. These are formed in a V_n with the help of the numerical tensor densities $\epsilon^{\mu_1\mu_2\ldots\mu_n}$ and $\epsilon_{\mu_1\mu_2\ldots\mu_n}$ [(**3**) 25]. How to determine their Hamiltonian derivatives is best illustrated by a simple example. Let $n=2$ and suppose that the Hamiltonian derivative of

$$\mathfrak{K}=\sqrt{(-\epsilon^{\mu\sigma}\epsilon^{\nu\rho}G_{\mu\nu}G_{\sigma\rho})} \quad (3.2)$$

be required. Raising the subscripts ν and ρ, we get

$$\mathfrak{K}=\sqrt{(-\epsilon^{\mu\sigma}\epsilon^{\nu\rho}g_{\nu\alpha}g_{\rho\beta}G^{\alpha}_{\mu}G^{\beta}_{\sigma})}. \quad (3.3)$$

But [(**3**) 5, (3.43)] $$\epsilon^{\nu\rho}g_{\nu\alpha}g_{\rho\beta}=\epsilon_{\alpha\beta}g. \quad (3.31)$$

Hence $$\mathfrak{K}=\gamma\sqrt{(\epsilon^{\mu\sigma}\epsilon_{\alpha\beta}G^{\alpha}_{\mu}G^{\beta}_{\sigma})}=\gamma\sqrt{(\delta^{\mu\sigma}_{\alpha\beta}G^{\alpha}_{\mu}G^{\beta}_{\sigma})},$$

where $\delta^{\mu\sigma}_{\alpha\beta}$ is a generalized Kronecker delta. Therefore

$$K=\mathfrak{K}/\gamma=\sqrt{(G^2-G_{\sigma\rho}G^{\sigma\rho})}, \quad (3.4)$$

which is of the form (3.11). Other relative invariants may be dealt with in an exactly similar way.

4. Consider a small variation of K in which the $g^{\mu\nu}$ are not varied. From (3.1) we have, since $\Gamma^{\mu_1\nu_1\dots\rho_m}$ involves only the $g^{\mu\nu}$,

$$\delta K = \sum_{r=1}^{m} \Omega^{\mu\nu\sigma\rho}_{(r)}\, \delta B_{\mu\nu\sigma\rho}, \tag{4.1}$$

where
$$\Omega^{\mu\nu\sigma\rho}_{(r)} = \Gamma^{\mu_1\nu_1\dots\rho_{r-1}\mu\nu\sigma\rho\mu_{r+1}\dots\rho_m} \prod_{s=1}^{m}{}' B_{\mu_s\nu_s\sigma_s\rho_s}, \tag{4.2}$$

the dash indicating that, in forming the product, the term $s = r$ is to be omitted. But

$$B_{\mu\nu\sigma\rho} = \tfrac{1}{2}(g_{\mu\nu,\sigma\rho}+g_{\sigma\rho,\mu\nu}-g_{\mu\sigma,\nu\rho}-g_{\nu\rho,\mu\sigma})+\{\mu\nu,\alpha\}[\sigma\rho,\alpha]-\{\mu\sigma,\alpha\}[\nu\rho,\alpha], \tag{4.3}$$

so that it is at once apparent from (4.1) that *all* the terms of $\partial K/\partial g_{\mu\nu,\alpha}$ have three-index symbols as factors. Hence, by (2.8), all the terms of $Y^{\mu\nu\alpha}$ have three-index symbols as factors. But $Y^{\mu\nu\alpha}$ is a tensor; and therefore it must be a zero tensor. Hence

$$\frac{\partial K}{\partial g_{\mu\nu,\alpha}} = -2\{\sigma\rho,\mu\}Z^{\nu\sigma\rho\alpha}-2\{\sigma\rho,\nu\}Z^{\mu\sigma\rho\alpha}-\{\sigma\rho,\alpha\}Z^{\mu\nu\sigma\rho}. \tag{4.4}$$

The coefficient of θ'' is accordingly simply

$$K_{;\alpha\beta}\, Z^{\mu\nu\alpha\beta}+2K_{;\alpha}\, Z^{\mu\nu\alpha\beta}{}_{;\beta}. \tag{4.5}$$

5. In re-expressing $P^{\mu\nu}$ as given by (2.3) it is unnecessary to write down explicitly those terms which are multiplied by three-index symbols, for such terms cannot by themselves form the components of a tensor [(**1**) 79], and we shall therefore represent them merely by dots. (This procedure is equivalent to using geodesic coordinates.)

Now, by (2.61) and (2.62), we have

$$\left.\begin{aligned} \gamma_{,\beta} &= \dots \\ \gamma_{,\alpha\beta} &= \gamma\{\beta\lambda,\lambda\}_{,\alpha}+\dots \end{aligned}\right\}, \tag{5.1}$$

and

$$Z^{\mu\nu\alpha\beta}{}_{,\alpha\beta} = Z^{\mu\nu\alpha\beta}{}_{,\beta\alpha} = (Z^{\mu\nu\alpha\beta}{}_{;\beta})_{,\alpha}-\{\sigma\beta,\mu\}_{,\alpha}\, Z^{\sigma\nu\alpha\beta}-\{\sigma\beta,\nu\}_{,\alpha}\, Z^{\mu\sigma\alpha\beta}- \\ -\{\sigma\beta,\alpha\}_{,\alpha}\, Z^{\mu\nu\sigma\beta}-\{\sigma\beta,\beta\}_{,\alpha}\, Z^{\mu\nu\alpha\sigma}+\dots, \tag{5.21}$$

$$(Z^{\mu\nu\alpha\beta}{}_{;\beta})_{,\alpha} = Z^{\mu\nu\alpha\beta}{}_{;\alpha\beta}+\dots.$$

Also write $\left[\partial K/\partial g_{\mu\nu}\right]$ for the 'tensor part' of $\partial K/\partial g_{\mu\nu}$, i.e. for those terms of $\partial K/\partial g_{\mu\nu}$ which form the components of a tensor, so that

$$\frac{\partial K}{\partial g_{\mu\nu}} = \left[\frac{\partial K}{\partial g_{\mu\nu}}\right]+\dots. \tag{5.22}$$

Using (2.4), (2.41) and substituting (5.1), (5.21) and (5.22) in (2.3) we reduce the latter to

$$P^{\mu\nu} = \tfrac{1}{2}g^{\mu\nu}K + \left[\frac{\partial K}{\partial g_{\mu\nu}}\right] + Z^{\mu\nu\alpha\beta}{}_{;\alpha\beta} + \{\sigma\rho,\mu\}_{,\alpha}\, Z^{\nu\sigma\rho\alpha} + \{\sigma\rho,\nu\}_{,\alpha}\, Z^{\mu\sigma\rho\alpha} + \ldots . \tag{5.3}$$

6. From (3.1) we see, as in § 7, that $Z^{\mu\nu\alpha\beta}$ arises from an expression of the form

$$z^{\mu\nu\sigma\rho}(\delta g_{\mu\nu,\sigma\rho} + \delta g_{\sigma\rho,\mu\nu} - \delta g_{\mu\sigma,\nu\rho} - \delta g_{\nu\rho,\mu\sigma}), \tag{6.1}$$

where $z^{\mu\nu\sigma\rho}$ is a tensor.

On changing dummy indices (6.1) becomes

$$(z^{\mu\nu\sigma\rho} + z^{\sigma\rho\mu\nu} - z^{\mu\sigma\nu\rho} - z^{\rho\nu\sigma\mu})\,\delta g_{\mu\nu,\sigma\rho} = \zeta^{\mu\nu\sigma\rho}\,\delta g_{\mu\nu,\sigma\rho} \quad \text{(say)}. \tag{6.11}$$

$Z^{\mu\nu\sigma\rho}$ is therefore the part of $\zeta^{\mu\nu\sigma\rho}$ which is symmetrical in μ, ν and in σ, ρ, i.e.

$$Z^{\mu\nu\sigma\rho} = \tfrac{1}{4}(\zeta^{\mu\nu\sigma\rho} + \zeta^{\nu\mu\sigma\rho} + \zeta^{\mu\nu\rho\sigma} + \zeta^{\nu\mu\rho\sigma}). \tag{6.2}$$

From (6.11) and (6.2) it may be verified that, in addition to (2.41), $Z^{\mu\nu\sigma\rho}$ satisfies the cyclic identity

$$Z^{\mu\nu\sigma\rho} + Z^{\mu\sigma\rho\nu} + Z^{\mu\rho\nu\sigma} = 0. \tag{6.3}$$

Hence $$\{\sigma\rho,\mu\}_{,\alpha}(Z^{\nu\sigma\rho\alpha} + Z^{\nu\rho\sigma\alpha} + Z^{\nu\alpha\sigma\rho}) = 0,$$

i.e., by (2.41),

$$\begin{aligned} 2\{\sigma\rho,\mu\}_{,\alpha}\, Z^{\nu\sigma\rho\alpha} &= -Z^{\nu\alpha\sigma\rho}\{\sigma\rho,\mu\}_{,\alpha} \\ &= -Z^{\nu\alpha\sigma\rho}(B_{\sigma\alpha\rho\cdot}{}^{\mu} + \{\sigma\alpha,\mu\}_{,\rho}) + \ldots \\ &= +Z^{\nu\alpha\sigma\rho}B^{\mu}{}_{\cdot\alpha\rho\sigma} - Z^{\nu\rho\sigma\alpha}\{\sigma\rho,\mu\}_{,\alpha} + \ldots \\ &= Z^{\nu\alpha\sigma\rho}B^{\mu}{}_{\cdot\alpha\sigma\rho} - Z^{\nu\sigma\rho\alpha}\{\sigma\rho,\mu\}_{,\alpha} + \ldots . \end{aligned}$$

Hence $$\{\sigma\rho,\mu\}_{,\alpha}\, Z^{\nu\sigma\rho\alpha} = \tfrac{1}{3}Z^{\nu\alpha\sigma\rho}B^{\mu}{}_{\cdot\alpha\sigma\rho} + \ldots . \tag{6.31}$$

Since $P^{\mu\nu}$ is a tensor, we have therefore from (5.3) that

$$P^{\mu\nu} = Z^{\mu\nu\sigma\rho}{}_{;\sigma\rho} + \left[\frac{\partial K}{\partial g_{\mu\nu}}\right] + \tfrac{1}{2}g^{\mu\nu}K + \tfrac{1}{3}(Z^{\nu\alpha\sigma\rho}B^{\mu}{}_{\cdot\alpha\sigma\rho} + Z^{\mu\alpha\sigma\rho}B^{\nu}{}_{\cdot\alpha\sigma\rho}). \tag{6.4}$$

The last part of (6.4) may be rewritten in a useful form as follows.

Consider the tensor $t^{\mu}_{\lambda} = Z^{\mu\nu\sigma\rho}B_{\lambda\nu\sigma\rho}$. By (6.2) this may be written, after some changes in dummy indices, as

$$\begin{aligned} t^{\mu}_{\lambda} = \tfrac{1}{4}[&(z^{\mu\nu\sigma\rho} + z^{\sigma\rho\mu\nu})(B_{\lambda\nu\sigma\rho} - B_{\lambda\sigma\nu\rho} + B_{\lambda\nu\rho\sigma} - B_{\lambda\sigma\rho\nu}) + \\ &+ (z^{\nu\mu\sigma\rho} + z^{\sigma\rho\nu\mu})(B_{\lambda\nu\sigma\rho} - B_{\lambda\rho\nu\sigma} + B_{\lambda\nu\rho\sigma} - B_{\lambda\rho\sigma\nu})], \end{aligned}$$

from which, in view of the symmetry properties of the curvature

tensor, and after some further changes in dummy indices, we find that

$$t^{\mu}_{\lambda} = \tfrac{3}{4} B_{\lambda\nu\sigma\rho}(z^{\mu\nu\sigma\rho} + z^{\nu\mu\rho\sigma} + z^{\sigma\rho\mu\nu} + z^{\rho\sigma\nu\mu}). \tag{6.5}$$

Therefore

$$P^{\mu\nu} = Z^{\mu\nu\sigma\rho}{}_{;\sigma\rho} + \left[\frac{\partial K}{\partial g_{\mu\nu}}\right] + \tfrac{1}{2} g^{\mu\nu} K + \\ + \tfrac{1}{4}\{B^{\nu}_{\cdot\alpha\sigma\rho}(z^{\mu\alpha\sigma\rho} + z^{\alpha\mu\rho\sigma} + z^{\sigma\rho\mu\alpha} + z^{\rho\sigma\alpha\mu}) + \\ + B^{\mu}_{\cdot\alpha\sigma\rho}(z^{\nu\alpha\sigma\rho} + z^{\alpha\nu\rho\sigma} + z^{\sigma\rho\nu\alpha} + z^{\rho\sigma\alpha\nu})\}. \tag{6.6}$$

7.(*a*) If we examine (3.1), we may think of K written formally as the product of m of the components of the curvature tensor with all the indices paired so that to every superscript there corresponds one and only one subscript. Then corresponding to each such pair of indices there is one factor $g^{\sigma\rho}$ in $\Gamma^{\mu_1\nu_1\ldots\rho_m}$. In forming δK this contributes $-g^{\sigma\mu}g^{\rho\nu}\,\delta g_{\mu\nu}$. So that, if in turn each superscript is replaced by μ, whilst the corresponding subscript is first raised and then replaced by ν, and all $2m$ tensors so obtained are added, then the symmetrical part of the sum is $-[\partial K/\partial g_{\mu\nu}]$.

(*b*) Comparing (4.1) and (6.1) we see at once that

$$z^{\mu\nu\sigma\rho} = \tfrac{1}{2} \sum_{r=1}^{m} \Omega^{\mu\nu\sigma\rho}_{(r)}. \tag{7.1}$$

Hence the first four of the last eight terms of (6.6), say $q^{\mu\nu}$, are

$$q^{\mu\nu} = \tfrac{1}{8} \sum_{r=1}^{m} (\Omega^{\mu\alpha\sigma\rho}_{(r)} B^{\nu}_{\cdot\alpha\sigma\rho} + \Omega^{\alpha\mu\sigma\rho}_{(r)} B_{\alpha\cdot\sigma\rho}{}^{\nu} + \Omega^{\sigma\rho\mu\alpha}_{(r)} B_{\sigma\rho\cdot\alpha}{}^{\nu} + \Omega^{\sigma\rho\alpha\mu}_{(r)} B_{\sigma\rho\alpha\cdot}{}^{\nu}). \tag{7.2}$$

Inspection of (7.2) shows at once that $-4q^{\mu\nu}$ is formed by exactly the same process as that described in § 7(*a*), if we notice that each component of the curvature tensor is now taken twice over, which cancels the factor $\frac{1}{2}$ on the right-hand side of (7.1). Now, where a superscript was replaced by μ the first time, it is replaced by ν the second time. It follows that $q^{\mu\nu}$ is symmetrical, so that

$$q^{\mu\nu} = q^{\nu\mu} = -\frac{1}{4}\left[\frac{\partial K}{\partial g_{\mu\nu}}\right] = +\tfrac{1}{3} Z^{\nu\alpha\sigma\rho} B^{\mu}_{\cdot\alpha\sigma\rho}. \tag{7.3}$$

Hence (6.4) becomes

$$P^{\mu\nu} = Z^{\mu\nu\sigma\rho}{}_{;\sigma\rho} - \tfrac{2}{3} Z^{\nu\alpha\sigma\rho} B^{\mu}_{\cdot\alpha\sigma\rho} + \tfrac{1}{2} g^{\mu\nu} K, \tag{7.4}$$

which establishes (1.1) for the case in which K has the form (3.1).

A simple formula may be derived in this case for the scalar

invariant $P = g_{\mu\nu} P^{\mu\nu}$. From the law of formation of $[\partial K/\partial g_{\mu\nu}]$ given in § 7(*a*) it follows at once that

$$g_{\mu\nu}\left[\frac{\partial K}{\partial g_{\mu\nu}}\right] = -2mK. \tag{7.5}$$

Hence, with (7.3), (7.4) easily yields

$$P = (g_{\mu\nu} Z^{\mu\nu\sigma\rho})_{;\sigma\rho} + (\tfrac{1}{2}n - m)K. \tag{7.6}$$

When $m = \frac{1}{2}n$, we see that P has necessarily the form of a divergence.

8. I now show that (7.4) holds in the general case when K has the form (3.11). Write

$$P_i^{\mu\nu} = \frac{hK_i}{hg_{\mu\nu}}, \quad \text{and} \quad f_i = \frac{\partial F}{\partial K_i}, \qquad f_{ij} = \frac{\partial^2 F}{\partial K_i \partial K_j}, \quad \text{etc.}$$

Also let

$$Z_i^{\mu\nu\sigma\rho} = \frac{\partial K_i}{\partial g_{\mu\nu,\sigma\rho}}.$$

Then, using the summation convention also for roman indices, it is not difficult to show after the manner of §§ 2, 4 that

$$P^{\mu\nu} = \frac{hK}{hg_{\mu\nu}} = \tfrac{1}{2}g^{\mu\nu}F + f_i(P_i^{\mu\nu} - \tfrac{1}{2}g^{\mu\nu}K_i) +$$

$$+ f_{ij}[(K_j)_{;\sigma\rho} Z_i^{\mu\nu\sigma\rho} + 2(K_j)_{;\sigma} Z_i^{\mu\nu\sigma}{}_{;\rho}] + f_{ijk}(K_j)_{;\sigma}(K_k)_{;\rho} Z_i^{\mu\nu\sigma\rho}. \tag{8.1}$$

Now consider (7.4) with $K = F$. We have

$$Z^{\mu\nu\sigma\rho} = f_i Z_i^{\mu\nu\sigma\rho}.$$

Therefore

$$P^{\mu\nu} = (f_i Z_i^{\mu\nu\sigma\rho})_{;\sigma\rho} - \tfrac{2}{3}B^{\mu}_{\cdot\,\alpha\sigma\rho} f_i Z_i^{\nu\alpha\sigma\rho} + \tfrac{1}{2}g^{\mu\nu}F. \tag{8.2}$$

But

$$(f_i Z_i^{\mu\nu\sigma\rho})_{;\sigma\rho} = f_i Z_i^{\mu\nu\sigma\rho}{}_{;\sigma} + f_{ij}(K_j)_{;\sigma} Z_i^{\mu\nu\sigma\rho}, \quad \text{etc.}$$

Hence comparing (8.2) and (8.1) we see that (7.4) applies when K has the general form (3.11). This completes the proof.

9. It is scarcely worth while verifying in full detail that the divergence of (7.4) vanishes identically. I shall merely outline a proof for the case in which K is of the form (3.1). We have

$$P^{\mu\nu}{}_{;\nu} = Z^{\mu\nu\sigma\rho}{}_{;\sigma\rho\nu} - \tfrac{2}{3}(Z^{\nu\alpha\sigma\rho}B^{\mu}_{\cdot\,\alpha\sigma\rho})_{;\nu} + \tfrac{1}{2}K^{;\mu}. \tag{9.1}$$

If we add to $Z^{\mu\nu\sigma\rho}{}_{;\sigma\rho\nu}$ the expressions $Z^{\mu\nu\sigma\rho}{}_{;\sigma\nu\rho}$ and $Z^{\mu\nu\sigma\rho}{}_{;\nu\sigma\rho}$, both of which may be re-expressed in terms of $Z^{\mu\nu\sigma\rho}{}_{;\sigma\rho\nu}$ by means of Ricci's identity, we obtain

$$3Z^{\mu\nu\sigma\rho}{}_{;\sigma\rho\nu} = 2B^{\mu}_{\cdot\,\rho\nu\alpha} Z^{\alpha\nu\sigma\rho}{}_{;\sigma}, \tag{9.2}$$

all other terms vanishing on account of the symmetry properties of $Z^{\mu\nu\sigma\rho}$ and $B_{\mu\nu\sigma\rho}$, and of $q^{\mu\nu}$. For the same reason, when we substitute (9.2) in (9.1), the latter can be reduced to

$$P^{\mu\nu}{}_{;\nu} = -\tfrac{2}{3}B^{\mu}{}_{.\alpha\sigma\rho;\nu}\,Z^{\nu\alpha\sigma\rho} + \tfrac{1}{2}K^{;\mu}. \tag{9.3}$$

From (3.1), (4.1) and (7.1) we have at once that

$$\tfrac{1}{2}K^{;\mu} = z^{\beta\alpha\sigma\rho}B_{\beta\alpha\sigma\rho}{}^{;\mu}. \tag{9.4}$$

If the first term of (9.3) be now rewritten in terms of $z^{\mu\nu\sigma\rho}$, which can be done by inspection of (6.4) and (6.6), it is not difficult to show by applying the Bianchi identity to the resulting expression that it just cancels the second term, so that

$$P^{\mu\nu}{}_{;\nu} \equiv 0. \tag{9.5}$$

10. I shall now consider a few examples of specific invariants K.

(i) $$K = G \quad (= g^{\mu\nu}g^{\sigma\rho}B_{\mu\nu\sigma\rho}). \tag{10.1}$$

By inspection $z^{\mu\nu\sigma\rho} = \tfrac{1}{2}g^{\mu\nu}g^{\sigma\rho}$, giving

$$Z^{\mu\nu\sigma\rho} = (g^{\mu\nu}g^{\sigma\rho} - \tfrac{1}{2}g^{\mu\sigma}g^{\nu\rho} - \tfrac{1}{2}g^{\mu\rho}g^{\nu\sigma}).$$

Therefore, at once, since here $Z^{\mu\nu\sigma\rho}$ vanishes,

$$\frac{hG}{hg_{\mu\nu}} = -G^{\mu\nu} + \tfrac{1}{2}g^{\mu\nu}G. \tag{10.2}$$

(ii) $$K = G^2 = K''' \quad \text{(say)}. \tag{10.3}$$

It is easiest to use (8.2) with $F = K_1^2$ and $K_1 = G$. It follows immediately that

$$P'^{\mu\nu} = \frac{hG^2}{hg_{\mu\nu}} = -2\{G^{;\mu\nu} + GG^{\mu\nu} - g^{\mu\nu}(\Box G + \tfrac{1}{4}G^2)\}. \tag{10.4}$$

[The symbol $\Box$ before any expression (...) represents the operation $g^{\sigma\rho}(\ldots)_{;\sigma\rho}$.]

(iii) $$K = G_{\mu\nu}G^{\mu\nu} = K' \quad \text{(say)}. \tag{10.5}$$

For a small variation of the $g_{\mu\nu,\sigma\rho}$ alone

$$\delta K' = 2G^{\mu\nu}\,\delta G_{\mu\nu} = 2G^{\mu\nu}g^{\sigma\rho}\,\delta B_{\mu\nu\sigma\rho},$$

so that $z^{\mu\nu\sigma\rho} = g^{\sigma\rho}G^{\mu\nu}$, and

$$Z^{\mu\nu\sigma\rho} = g^{\mu\nu}G^{\sigma\rho} + g^{\sigma\rho}G^{\mu\nu} - \tfrac{1}{2}g^{\mu\sigma}G^{\nu\rho} - \tfrac{1}{2}g^{\nu\sigma}G^{\mu\rho} - \tfrac{1}{2}g^{\mu\rho}G^{\nu\sigma} - \tfrac{1}{2}g^{\nu\rho}G^{\mu\sigma}.$$

Using the identities of Ricci and Bianchi we find

$$Z^{\mu\nu\sigma\rho}{}_{;\sigma\rho} = \Box G^{\mu\nu} - G^{;\mu\nu} - B^{\mu\nu\sigma\rho}_{\ldots\ldots}G_{\sigma\rho} + G^{\mu\rho}G^{\nu}_{\rho} + \tfrac{1}{2}g^{\mu\nu}\Box G.$$

Also $$Z^{\mu\alpha\sigma\rho}B^{\nu}{}_{.\alpha\sigma\rho} = \tfrac{3}{2}(B^{\mu\nu\sigma\rho}_{\ldots\ldots}G_{\sigma\rho} + G^{\mu\rho}G^{\nu}_{\rho}).$$

Therefore

$$P'^{\mu\nu} = \frac{hK'}{hg_{\mu\nu}} = \Box G^{\mu\nu} - G^{;\mu\nu} - 2B^{\mu\nu\sigma\rho}_{\cdot\cdot\cdot\cdot} G_{\sigma\rho} + \tfrac{1}{2} g^{\mu\nu}(\Box G + G_{\sigma\rho} G^{\sigma\rho}). \tag{10.6}$$

(iv) $$K = B_{\mu\nu\sigma\rho} B^{\mu\nu\sigma\rho}_{\cdot\cdot\cdot\cdot} = K'' \quad \text{(say)}. \tag{10.7}$$

By inspection, $z^{\mu\nu\sigma\rho} = B^{\mu\nu\sigma\rho}$, which, in view of the symmetry properties of $B_{\mu\nu\sigma\rho}$, gives

$$Z^{\mu\nu\sigma\rho} = 2(B^{\mu\nu\sigma\rho}_{\cdot\cdot\cdot\cdot} + B^{\mu\nu\rho\sigma}_{\cdot\cdot\cdot\cdot}).$$

Hence $$Z^{\mu\nu\sigma\rho}{}_{;\sigma\rho} = 4\Box G^{\mu\nu} - 2G^{;\mu\nu} - 4B^{\mu\nu\sigma\rho}_{\cdot\cdot\cdot\cdot} G_{\sigma\rho} + 4G^{\mu\sigma} G^{\nu}_{\sigma}.$$

Again, in view of the symmetry properties of the curvature tensor

$$Z^{\mu\alpha\sigma\rho} B^{\nu}_{\cdot\alpha\sigma\rho} = 2(B^{\mu\alpha\sigma\rho}_{\cdot\cdot\cdot\cdot} B^{\nu}_{\cdot\alpha\sigma\rho} + B^{\mu\alpha\sigma\rho}_{\cdot\cdot\cdot\cdot} B^{\nu}_{\cdot\alpha\rho\sigma}) = 3B^{\mu\alpha\sigma\rho}_{\cdot\cdot\cdot\cdot} B^{\nu}_{\cdot\alpha\sigma\rho}.$$

Therefore

$$P''^{\mu\nu} = \frac{hK''}{hg_{\mu\nu}}$$
$$= 4\{\Box G^{\mu\nu} - \tfrac{1}{2} G^{;\mu\nu} - B^{\mu\nu\sigma\rho}_{\cdot\cdot\cdot\cdot} G_{\sigma\rho} + G^{\mu\sigma} G^{\nu}_{\sigma} - \tfrac{1}{2} B^{\mu\alpha\sigma\rho}_{\cdot\cdot\cdot\cdot} B^{\nu}_{\cdot\alpha\sigma\rho} + \tfrac{1}{8} g^{\mu\nu} K''\}. \tag{10.8}$$

11. In the case of the three invariants K', K'', K''', we have $m = 2$, so that, according to (7.6), their contracted Hamiltonian derivatives P', P'', P''' must each have the form of a divergence in a V_4. In fact we find

$$\tfrac{1}{2} P' = \tfrac{1}{2} P'' = \tfrac{1}{6} P''' = \Box G. \tag{11.1}$$

It will be noticed that the terms involving the fourth derivatives of the $g_{\mu\nu}$ can be eliminated between the three Hamiltonian derivatives (10.4), (10.6), and (10.8). Thus, if we write

$$\left.\begin{aligned} L^{\mu\nu} &= B^{\mu\alpha\sigma\rho}_{\cdot\cdot\cdot\cdot} B^{\nu}_{\cdot\alpha\sigma\rho} + GG^{\mu\nu} - 2G^{\mu\sigma} G^{\nu}_{\sigma} - 2B^{\mu\nu\sigma\rho}_{\cdot\cdot\cdot\cdot} G_{\sigma\rho} \\ L &= g_{\mu\nu} L^{\mu\nu} = -4K' + K'' + K''' \end{aligned}\right\}, \tag{11.2}$$

we have $$-\frac{1}{2}\frac{hL}{hg_{\mu\nu}} = L^{\mu\nu} - \tfrac{1}{2} g^{\mu\nu} L. \tag{11.3}$$

Accordingly there exists in a V_n ($n > 4$, see below) a linear combination with constant coefficients of the three invariants K', K'', K''', namely L, such that its Hamiltonian derivative is a tensor of the *second* differential order. This naturally implies a considerable restriction on any common integrals which the three sets of differential equations

$$P'^{\mu\nu} = 0, \qquad P''^{\mu\nu} = 0, \qquad P'''^{\mu\nu} = 0 \tag{11.4}$$

may possess, for they must satisfy the set of *second*-order equations

$$L^{\mu\nu} = 0. \tag{11.5}$$

It has been shown by Lanczos (**2**) that $\hbar L/hg_{\mu\nu}$ vanishes identically in a V_4. The same result holds in a V_2 and a V_3.

REFERENCES

1. A. S. Eddington, *The Mathematical Theory of Relativity* (2nd ed. Cambridge, 1930).
2. C. Lanczos, *Annals of Math.* (2) **39** (1938), 842.
3. O. Veblen, *Invariants of Quadratic Differential Forms* (Cambridge Tract **24**, 1927).

ON THE EQUATION $a_1x_1^k+a_2x_2^k+\ldots+a_kx_k^k+b=0$ IN FINITE FIELDS

By ŠT. SCHWARZ (*Bratislava*)

[Received 5 November 1947]

IN this paper I prove the following theorem:*

THEOREM. *Suppose that* (i) $GF(p^n)$ *is a finite field of characteristic* p; (ii) $a_1, a_2,\ldots, a_k, b$ *are elements of the field* $GF(p^n)$; (iii) $k|(p-1)$, $ba_1a_2\ldots a_k \neq 0$, $k \geqslant 2$. *Then the equation*

$$a_1x_1^k+a_2x_2^k+\ldots+a_kx_k^k+b=0 \tag{1}$$

has at least one solution in which $x_1, x_2,\ldots, x_k \in GF(p^n)$.

Proof. Consider the expression

$$(a_1x_1^k+a_2x_2^k+\ldots+a_kx_k^k+b)^{p^n-1}. \tag{2}$$

In $GF(p^n)$ this expression is equal to 0 (*zero element*) or 1 (*unit element*) according as the bracketed expression in (2) is zero or any non-zero element of the field $GF(p^n)$. Consider further the sum

$$\sigma=\sum_{x_1=\xi}\sum_{x_2=\xi}\ldots\sum_{x_k=\xi}(a_1x_1^k+a_2x_2^k+\ldots+a_kx_k^k+b)^{p^n-1}, \tag{3}$$

where $x_1, x_2,\ldots, x_k$ run through all elements ξ of the field $GF(p^n)$. Suppose that there is a set of elements $a_1, a_2,\ldots, a_k, b$ with $ba_1a_2\ldots a_k \neq 0$ for which (1) has no solution. Then every summand in (3) taken with these $a_1, a_2,\ldots, a_k, b$ is equal to unity. We obtain summarily p^{kn} unities, and the element σ would therefore be equal to zero. If we can show that, for any $a_1, a_2,\ldots, a_k, b$, the number σ is different from zero, we shall have proved that some of the summands in (3) are equal to zero and therefore that the equation (1) has at least one solution in which $x_1, x_2,\ldots, x_k \in GF(p^n)$.

On the right of (3) we have

$$(a_1x_1^k+a_2x_2^k+\ldots+a_kx_k^k+b)^{p^n-1}$$

$$=\sum_{\lambda_1,\ldots,\lambda_k,\lambda_{k+1}}\frac{(p^n-1)!}{\lambda_1!\ldots\lambda_k!\,\lambda_{k+1}!}a_1^{\lambda_1}\ldots a_k^{\lambda_k}b^{\lambda_{k+1}}x_1^{k\lambda_1}\ldots x_k^{k\lambda_k}, \tag{4}$$

* This theorem contains as special cases the results due to L. Tornheim, 'Sums of nth powers in fields of prime characteristic', *Duke Math. J.* 4 (1938), 359–62, and L. Rédei, 'Zur Theorie der Gleichungen in end[illegible]en Körpern', *Acta Univ. Szeged*, 11 (1946), 68. See also Št. Schwarz, 'On Waring's problem for finite fields', pp. 123–8, above.

where the summation extends over the sets $\lambda_1,\ldots,\lambda_k,\lambda_{k+1}$ of non-negative integers satisfying

$$\lambda_1+\ldots+\lambda_k+\lambda_{k+1}=p^n-1. \tag{5}$$

We remark that the number $(p^n-1)!/\lambda_1!\ldots\lambda_k!\lambda_{k+1}!$ is an integer. In what follows I consider the residue of this number to modulus p. The number p may appear formally in the denominator of this fraction. Whenever we take this number to modulus p, it is naturally understood that such a factor p is first cancelled with p in the numerator, so that we have really an integer.

Substituting (4) in (3) we have

$$\sigma=\sum_{\lambda_1,\ldots,\lambda_k,\lambda_{k+1}}\sum_{x_1=\xi}\ldots\sum_{x_k=\xi}\frac{(p^n-1)!}{\lambda_1!\ldots\lambda_k!\lambda_{k+1}!}a_1^{\lambda_1}\ldots a_k^{\lambda_k}b^{\lambda_{k+1}}x_1^{k\lambda_1}\ldots x_k^{k\lambda_k}. \tag{6}$$

To establish the value of σ, I use the following well-known result:

If ξ runs through all p^n elements of $GF(p^n)$ then

$$\sum_{x=\xi}x^\lambda=\begin{cases}0 & (p^n-1\nmid\lambda),\\ -1 & (p^n-1\mid\lambda;\ \lambda>0).\end{cases} \tag{7}$$

It follows from this formula that those (and only those) of the summands in (6) will be different from zero in which all exponents $k\lambda_1$, $k\lambda_2,\ldots,k\lambda_k$ are divisible by p^n-1. Moreover, we must have

$$\lambda_1>0,\quad \lambda_2>0,\quad \ldots,\quad \lambda_k>0.$$

For, suppose that at least one of these numbers, say λ_1, is equal to zero. Then the summand corresponding to $(0,\lambda_2,\lambda_3,\ldots,\lambda_k,\lambda_{k+1})$ is of the form

$$\begin{aligned}&\sum_{x_1=\xi}\sum_{x_2=\xi}\ldots\sum_{x_k=\xi}\frac{(p^n-1)!}{\lambda_2!\ldots\lambda_k!\lambda_{k+1}!}a_2^{\lambda_2}\ldots a_k^{\lambda_k}b^{\lambda_{k+1}}x_2^{k\lambda_2}\ldots x_k^{k\lambda_k}\\ &\quad=\left\{\sum_{x_2=\xi}\ldots\sum_{x_k=\xi}\frac{(p^n-1)!}{\lambda_2!\ldots\lambda_k!\lambda_{k+1}!}a_2^{\lambda_2}\ldots a_k^{\lambda_k}b^{\lambda_{k+1}}x_2^{k\lambda_2}\ldots x_k^{k\lambda_k}\right\}\sum_{x_1=\xi}1\\ &\quad=\{\ldots\}p^n=0.\end{aligned}$$

That is, in summing through x_1 we have to add p^n times the expression in the bracket, so that the sum is equal to zero.

Thus to obtain the value of σ we may omit in (6) all summands except those for which the following relations hold:

$$k\lambda_i\equiv 0\ (\mathrm{mod}\ p^n-1),\qquad \lambda_i>0\quad (i=1,2,\ldots,k).$$

The last relations can be written in the form

$$\lambda_i=n_i\frac{p^n-1}{k},\quad \text{where}\quad n_i>0\quad (i=1,2,\ldots,k)$$

are positive integers. From the equation (5),

$$\sum_{i=1}^{k} \frac{n_i(p^n-1)}{k} + \lambda_{k+1} = p^n - 1,$$

we have $$n_i = 1 \quad (i = 1, 2, \ldots, k),$$

i.e. $$\lambda_1 = \lambda_2 = \ldots = \lambda_k = \frac{p^n-1}{k}, \qquad \lambda_{k+1} = 0.$$

Accordingly

$$\sigma = (a_1 a_2 \ldots a_k)^{(p^n-1)/k} \sum_{x_1=\xi} \ldots \sum_{x_k=\xi} \frac{(p^n-1)!}{[\{(p^n-1)/k\}!]^k} x_1^{p^n-1} \ldots x_k^{p^n-1}$$

and, after the summation with respect to (7),

$$\sigma = (-1)^k (a_1 a_2 \ldots a_k)^{(p^n-1)/k} \frac{(p^n-1)!}{[\{(p^n-1)/k\}!]^k}. \tag{8}$$

I shall show that $\sigma \neq 0$. It is sufficient to prove that the number $(p^n-1)!/[\{(p^n-1)/k\}!]^k$ is not divisible by p. Let us determine the highest power of the prime p that divides the factorials $(p^n-1)!$ and $[\{(p^n-1)/k\}!]^k$. I use the well-known

LEMMA. *If p^μ is the highest power of the prime p which divides $m!$ and if $[x]$ denotes the greatest integer $\leqslant x$, then*

$$\mu = \left[\frac{m}{p}\right] + \left[\frac{m}{p^2}\right] + \left[\frac{m}{p^3}\right] + \ldots = \frac{m-s}{p-1},$$

where $s = \alpha_0 + \alpha_1 + \ldots + \alpha_l$ is the sum of the digits of m to the base p, so that

$$m = \alpha_0 p^l + \alpha_1 p^{l-1} + \ldots + \alpha_l \quad (0 \leqslant \alpha_i < p).$$

We can write the number p^n-1 in the form

$$p^n - 1 = \alpha_1 p^{n-1} + \alpha_2 p^{n-2} + \ldots + \alpha_n,$$

where $\alpha_1 = \alpha_2 = \ldots = \alpha_n = p-1$. We obtain therefore for the highest power p^μ dividing $(p^n-1)!$

$$\mu = \frac{1}{p-1}\{p^n - 1 - n(p-1)\}.$$

Further let p^ν be the highest power of p dividing $[\{(p^n-1)/k\}!]^k$. Since $(p^n-1)/k$ written to the base p is of the form*

$$\frac{p^n-1}{k} = \alpha_1 p^{n-1} + \alpha_2 p^{n-2} + \ldots + \alpha_n,$$

* Here we essentially use the supposition $k \mid p-1$.

where $\alpha_1 = \alpha_2 = \ldots = \alpha_n = (p-1)/k$, we have

$$\nu = k\frac{1}{p-1}\left\{\frac{p^n-1}{k} - n\frac{p-1}{k}\right\}.$$

Consequently $\nu = \mu$. In (8) we may cancel p^μ in the numerator and denominator.

We are left with $\sigma \neq 0$, and the proof is complete.*

Finally, I am grateful to a referee on whose advice I removed an obscurity in the text.

* We can go farther, though it is irrelevant for our purpose, and simplify the relation (8). If p^μ is again the highest power of the prime p which divides $m!$, and $m = \alpha_0 p^l + \alpha_1 p^{l-1} + \ldots + \alpha_l$ $(0 \leqslant \alpha_i < p)$, it can be proved that

$$\frac{m!}{p^\mu} \equiv (-1)^\mu \alpha_0! \, \alpha_1! \ldots \alpha_l! \pmod{p}.$$

(See, for example, Stickelberger, *Math. Ann.* 37 (1890), 321; Dickson, *History of the Theory of Numbers* 1 (1919), 263.) We therefore have

$$\frac{(p^n-1)!}{p^\mu} \equiv (-1)^\mu\{(p-1)!\}^n, \quad \frac{\{(p^n-1)/k\}!}{p^{\mu/k}} \equiv (-1)^{\mu/k}[\{(p-1)/k\}!]^n \pmod{p}$$

and

$$\sigma = (-1)^{n+k}(a_1 a_2 \ldots a_k)^{(p^n-1)/k}[\{(p-1)/k\}!]^{-nk}.$$

COMPLEX FOURIER–BESSEL TRANSFORMS

By E. C. TITCHMARSH (*Oxford*)

[Received 17 November 1947]

1. THE Fourier-integral formula for a given arbitrary function $f(x)$ can be written in the form

$$f(x) = \frac{1}{\sqrt{(2\pi)}} \int_{-\infty}^{\infty} F(u)e^{-ixu}\,du, \tag{1.1}$$

where

$$F(u) = \frac{1}{\sqrt{(2\pi)}} \int_{-\infty}^{\infty} f(t)e^{iut}\,dt. \tag{1.2}$$

This is true, for example, if $f(t)$ is absolutely integrable over $(-\infty,\infty)$, and of bounded variation in the neighbourhood of $t = x$, $f(x)$ in (1.1) being replaced where necessary by $\frac{1}{2}\{f(x+0)+f(x-0)\}$.

If $f(t)$ is not absolutely integrable over $(-\infty,\infty)$, but $e^{-c|t|}f(t)$ is, for some positive c, we can replace (1.1) by*

$$f(x) = \frac{1}{\sqrt{(2\pi)}} \int_{ia-\infty}^{ia+\infty} F_+(z)e^{-ixz}\,dz + \frac{1}{\sqrt{(2\pi)}} \int_{ib-\infty}^{ib+\infty} F_-(z)e^{-ixz}\,dz, \tag{1.3}$$

where $a > c$, $b < -c$, and

$$F_+(z) = \frac{1}{\sqrt{(2\pi)}} \int_0^{\infty} f(t)e^{izt}\,dt, \qquad F_-(z) = \frac{1}{\sqrt{(2\pi)}} \int_{-\infty}^{0} f(t)e^{izt}\,dt. \tag{1.4}$$

The object of this paper is to obtain a corresponding extension of Hankel's formula

$$f(x) = \int_0^{\infty} J_\nu(xu)(xu)^{\frac{1}{2}}\,du \int_0^{\infty} J_\nu(ut)(ut)^{\frac{1}{2}}f(t)\,dt, \tag{1.5}$$

which holds if $f(t)$ satisfies the conditions for (1.1), (1.2), and $\nu \geqslant -\frac{1}{2}$. With this may be associated my formula†

$$f(x) = \int_0^{\infty} \mathbf{H}_\nu(xu)(xu)^{\frac{1}{2}}\,du \int_0^{\infty} Y_\nu(ut)(ut)^{\frac{1}{2}}f(t)\,dt, \tag{1.6}$$

* See my *Fourier Integrals*.

† E. C. Titchmarsh, *Proc. London Math. Soc.* **22** (1923), xxxiv–xxxv, and *Fourier Integrals*, § 8.4.

where Y_ν is the Bessel function of the second kind, $\mathbf{H}_\nu$ is Struve's function,* and $-\frac{1}{2} \leqslant \nu < \frac{1}{2}$. It is now seen that, if $-\frac{1}{2} \leqslant \nu < \frac{1}{2}$, these may be combined into the formula

$$f(x) = \tfrac{1}{2} \int_{ia-\infty}^{ia+\infty} \{J_\nu(xz) - i\mathbf{H}_\nu(xz)\}(xz)^{\frac{1}{2}} F(z)\, dz, \tag{1.7}$$

where $a > 0$, and

$$F(z) = \tfrac{1}{2} \int_0^\infty H_\nu^{(1)}(tz)(tz)^{\frac{1}{2}} f(t)\, dt. \tag{1.8}$$

Here $H_\nu^{(1)}(z) = J_\nu(z) + iY_\nu(z)$.

There are formulae similar to (1.3), (1.4) if $f(t)$ is defined over $(-\infty, \infty)$ instead of over $(0, \infty)$, and extensions to $\nu \geqslant \frac{1}{2}$, which will be obtained later.

2. Proofs of (1.5) are given in Watson, *Theory of Bessel Functions*, Ch. 14, and in *Fourier Integrals*, § 8.18. I shall begin by giving the corresponding proof for (1.6).

THEOREM. *Let* $-\frac{1}{2} \leqslant \nu < \frac{1}{2}$, *and let*

$$\int_0^1 t^{-\frac{1}{2}-\nu} |f(t)|\, dt, \qquad \int_1^\infty |f(t)|\, dt$$

be finite. Let $f(t)$ *be of bounded variation in the neighbourhood of* $t = x > 0$. *Then* (1.6) *holds,* $f(x)$ *being replaced by* $\frac{1}{2}\{f(x+0) + f(x-0)\}$, *and the u-integral denoting* $\lim_{\lambda\to\infty} \int_0^\lambda$.

Since $Y_\nu(ut)(ut)^{\frac{1}{2}}$ is bounded if $-\frac{1}{2} \leqslant \nu < \frac{1}{2}$, the inner integral in (1.6) is uniformly convergent over any finite interval of u. Hence, if $\delta > 0$,

$$I = \int_0^\lambda \mathbf{H}_\nu(xu)(xu)^{\frac{1}{2}}\, du \int_0^\infty Y_\nu(ut)(ut)^{\frac{1}{2}} f(t)\, dt$$

$$= \int_0^\lambda \left\{ \int_0^{x-\delta} + \int_{x-\delta}^{x} + \int_x^{x+\delta} + \int_{x+\delta}^\infty \right\} = I_1 + I_2 + I_3 + I_4 \text{ (say)},$$

and we may invert the orders of integration.

Now $\phi = \mathbf{H}_\nu(xu)$ satisfies

$$\frac{d^2\phi}{du^2} + \frac{1}{u}\frac{d\phi}{du} + \left(x^2 - \frac{\nu^2}{u^2}\right)\phi = \frac{(\frac{1}{2}u)^{\nu-1} x^{\nu+1}}{\Gamma(\nu+\frac{1}{2})\sqrt{\pi}},$$

* See Watson, *Theory of Bessel Functions*, § 10.4.

and $\psi = Y_\nu(tu)$ satisfies

$$\frac{d^2\psi}{du^2}+\frac{1}{u}\frac{d\psi}{du}+\left(t^2-\frac{\nu^2}{u^2}\right)\psi = 0.$$

Hence

$$\frac{d}{du}\{u(\phi\psi'-\phi'\psi)\}+(t^2-x^2)u\phi\psi = -\frac{u^\nu x^{\nu+1}\psi}{2^{\nu-1}\Gamma(\nu+\frac{1}{2})\sqrt{\pi}}.$$

It follows that

$$(x^2-t^2)\int_0^\lambda \mathbf{H}_\nu(xu)Y_\nu(tu)u\,du$$

$$= [tu\mathbf{H}_\nu(xu)Y'_\nu(tu)-xu\mathbf{H}'_\nu(xu)Y_\nu(tu)]_0^\lambda+\frac{x^{\nu+1}}{2^{\nu-1}\Gamma(\nu+\frac{1}{2})\sqrt{\pi}}\int_0^\lambda u^\nu Y_\nu(tu)\,du.$$

For small z,

$$\mathbf{H}_\nu(z) = O(z^{\nu+1}),\ \mathbf{H}'_\nu(z) = O(z^\nu),\ Y_\nu(z) = O(z^{-|\nu|}),\ Y'_\nu(z) = O(z^{-|\nu|-1}),$$

with logarithmic factors if $\nu = 0$. Hence the integrated term vanishes at the lower limit. Therefore

$$I_1 = x^{\frac{1}{2}}\lambda\mathbf{H}_\nu(x\lambda)\int_0^{x-\delta}\frac{Y'_\nu(t\lambda)}{x^2-t^2}t^{\frac{3}{2}}f(t)\,dt-x^{\frac{3}{2}}\lambda\mathbf{H}'_\nu(x\lambda)\int_0^{x-\delta}\frac{Y_\nu(t\lambda)}{x^2-t^2}t^{\frac{1}{2}}f(t)\,dt+$$

$$+\frac{x^{\nu+\frac{3}{2}}}{2^{\nu-1}\Gamma(\nu+\frac{1}{2})\sqrt{\pi}}\int_0^{x-\delta}\frac{t^{\frac{1}{2}}f(t)}{x^2-t^2}\,dt\int_0^\lambda u^\nu Y_\nu(tu)\,du$$

$$= I_{1,1}+I_{1,2}+I_{1,3}\ \text{(say)}.$$

Now,* when $|\arg z| < \pi$ and $|z|$ is large,

$$\mathbf{H}_\nu(z) = Y_\nu(z)+\sum_{m=0}^{p-1}\frac{b_{m,\nu}}{z^{2m-\nu+1}}+O(z^{\nu-2p-1}), \tag{2.1}$$

where

$$b_{m,\nu} = \frac{2^{2m-\nu+1}\Gamma(m+\frac{1}{2})}{\pi\Gamma(\nu+\frac{1}{2}-m)};$$

(here the case where $p = 0$ and the sum vanishes is needed); and

$$Y_\nu(z) = \left(\frac{2}{\pi z}\right)^{\frac{1}{2}}\sin(z-\tfrac{1}{2}\nu\pi-\tfrac{1}{4}\pi)+O(z^{-\frac{3}{2}}).$$

The formulae obtained by formal differentiation are also valid.

* Watson, ibid., § 10.42.

Hence

$$I_{1,2} = O(\lambda^{\frac{1}{2}}) \int_0^{1/\lambda} (t\lambda)^{-|\nu|} t^{\frac{1}{2}} |f(t)|\, dt +$$

$$+ O\left\{ \int_{1/\lambda}^{x-\delta} \sin(t\lambda - \tfrac{1}{2}\nu\pi - \tfrac{1}{4}\pi) \frac{f(t)}{x^2 - t^2}\, dt \right\} + O\left\{ \frac{1}{\lambda} \int_{1/\lambda}^{x-\delta} \frac{|f(t)|}{t}\, dt \right\}.$$

The first term is

$$O\left\{ \int_0^{1/\lambda} (t\lambda)^{\frac{1}{2}-|\nu|} |f(t)|\, dt \right\} = O\left\{ \int_0^{1/\lambda} |f(t)|\, dt \right\} = o(1),$$

the second term is $o(1)$, by the Riemann–Lebesgue theorem, and the third term is

$$O\left\{ \int_0^{\lambda^{-\frac{1}{2}}} |f(t)|\, dt \right\} + O\left\{ \lambda^{-\frac{1}{2}} \int_{\lambda^{-\frac{1}{2}}}^{x-\delta} |f(t)|\, dt \right\} = o(1).$$

Hence $I_{1,2} \to 0$, and similarly $I_{1,1} \to 0$.

Now, if $t\lambda \leqslant 1$,

$$\int_0^{\lambda} u^{\nu} Y_{\nu}(tu)\, du = O\left\{ \int_0^{\lambda} u^{\nu} (tu)^{-|\nu|}\, du \right\} = O(t^{-|\nu|} \lambda^{1+\nu-|\nu|}).$$

Also*

$$\int_0^{\infty} u^{\nu} Y_{\nu}(u)\, du = 0 \quad (-\tfrac{1}{2} \leqslant \nu < \tfrac{1}{2}).$$

Hence, if $t\lambda > 1$,

$$\int_0^{\lambda} u^{\nu} Y_{\nu}(tu)\, du = -\int_{\lambda}^{\infty} u^{\nu} Y_{\nu}(tu)\, du = -t^{-\nu-1} \int_{t\lambda}^{\infty} s^{\nu} Y_{\nu}(s)\, ds$$

$$= O\{t^{-\nu-1}(t\lambda)^{\nu-\frac{1}{2}}\} = O(t^{-\frac{3}{2}} \lambda^{\nu-\frac{1}{2}}).$$

Consequently

$$I_{1,3} = O\left\{ \lambda^{1+\nu-|\nu|} \int_0^{1/\lambda} t^{\frac{1}{2}-|\nu|} |f(t)|\, dt \right\} + O\left\{ \lambda^{\nu-\frac{1}{2}} \int_{1/\lambda}^{x-\delta} t^{-1} |f(t)|\, dt \right\}$$

$$= O\left\{ \int_0^{1/\lambda} (t\lambda)^{1+\nu-|\nu|} t^{-\frac{1}{2}-\nu} |f(t)|\, dt \right\} + O\left\{ \int_{1/\lambda}^{x-\delta} (t\lambda)^{\nu-\frac{1}{2}} t^{-\frac{1}{2}-\nu} |f(t)|\, dt \right\}$$

$$= O\left\{ \int_0^{1/\lambda} t^{-\frac{1}{2}-\nu} |f(t)|\, dt \right\} + O\left\{ \int_{1/\lambda}^{\lambda^{-\frac{1}{2}}} t^{-\frac{1}{2}-\nu} |f(t)|\, dt \right\} +$$

$$+ O\left\{ \lambda^{\frac{1}{2}\nu-\frac{1}{4}} \int_{\lambda^{-\frac{1}{2}}}^{x-\delta} t^{-\frac{1}{4}-\nu} |f(t)|\, dt \right\} = o(1).$$

Hence $I_1 \to 0$. By a similar but simpler argument, $I_4 \to 0$.

* Watson, *Theory of Bessel Functions*, § 13.24 (5).

3. For the discussion of I_2 and I_3 we require the formula

$$\int_0^\infty \mathbf{H}_\nu(xu)Y_{\nu+1}(\xi u)\,du = \begin{cases} -x^\nu/\xi^{\nu+1} & (x>\xi), \\ -1/2\xi & (x=\xi), \\ 0 & (x<\xi). \end{cases} \tag{3.1}$$

This follows from the Mellin-transform formula*

$$\int_0^\infty f(xu)g(\xi u)\,du = \frac{1}{2\pi i}\int_{k-i\infty}^{k+i\infty} \mathfrak{F}(s)\mathfrak{G}(1-s)x^{-s}\xi^{s-1}\,ds$$

with

$$f(u) = \mathbf{H}_\nu(u), \qquad \mathfrak{F}(s) = \frac{2^{s-1}\Gamma(\frac{1}{2}s+\frac{1}{2}\nu)\tan\frac{1}{2}(s+\nu)\pi}{\Gamma(\frac{1}{2}\nu-\frac{1}{2}s+1)},$$

$$g(u) = Y_{\nu+1}(u),$$

$$\mathfrak{G}(s) = -\frac{2^{s-1}}{\pi}\Gamma(\tfrac{1}{2}s+\tfrac{1}{2}\nu+\tfrac{1}{2})\Gamma(\tfrac{1}{2}s-\tfrac{1}{2}\nu-\tfrac{1}{2})\cos(\tfrac{1}{2}s-\tfrac{1}{2}\nu-\tfrac{1}{2})\pi,$$

and $-\nu-1 < k < -\nu$. After some reduction the right-hand side becomes

$$\frac{1}{2\pi i\xi}\int_{k-i\infty}^{k+i\infty} \frac{(\xi/x)^s}{s+\nu}\,ds,$$

and the result follows. The process can be justified as in the proof of (7.11.15) in *Fourier Integrals*.

We can suppose δ so small that $f(t)$ is of bounded variation in $(x-\delta, x+\delta)$. Then in $(x, x+\delta)$ we can write

$$t^{-\nu-\frac{1}{2}}f(t) = x^{-\nu-\frac{1}{2}}f(x+0)+\chi_1(t)-\chi_2(t),$$

where χ_1 and χ_2 are positive, increasing, and less than ϵ. Then

$$I_3 = x^{\frac{1}{2}}\int_0^\lambda \mathbf{H}_\nu(xu)u\,du\int_x^{x+\delta} Y_\nu(ut)t^{\nu+1}\{x^{-\nu-\frac{1}{2}}f(x+0)+\chi_1(t)-\chi_2(t)\}\,dt.$$

The first term in the bracket contributes

$$x^{-\nu}f(x+0)\int_0^\lambda \mathbf{H}_\nu(xu)[Y_{\nu+1}\{(x+\delta)u\}(x+\delta)^{\nu+1}-Y_{\nu+1}(xu)x^{\nu+1}]\,du \to \tfrac{1}{2}f(x+0),$$

by (3.1). By the second mean-value theorem, the second term contributes

$$x^{\frac{1}{2}}\chi_1(x+\delta)\int_\xi^{x+\delta} t^{\nu+1}\,dt\int_0^\lambda \mathbf{H}_\nu(xu)Y_\nu(ut)u\,du$$

$$= x^{\frac{1}{2}}\chi_1(x+\delta)\int_0^\lambda \mathbf{H}_\nu(xu)[Y_{\nu+1}\{(x+\delta)u\}(x+\delta)^{\nu+1}-Y_{\nu+1}(\xi u)\xi^{\nu+1}]\,du \qquad (x<\xi<x+\delta).$$

* *Fourier Integrals* (2.1.23), (7.9.13), and (7.9.7).

It is easily seen from the asymptotic expansions that this integral is bounded for all λ, and $x < \xi < x+\delta$. The whole term is therefore $O(\epsilon)$. A similar argument applies to the term involving χ_2. Hence $I_3 \to \frac{1}{2}f(x+0)$. Similarly $I_2 \to \frac{1}{2}f(x-0)$, and the theorem follows.

4. We can now prove the simplest form of the complex transform theorem.

THEOREM. *Let* $-\frac{1}{2} \leqslant \nu < \frac{1}{2}$, $a > 0$, *let*

$$\int_0^1 t^{-\frac{1}{2}-\nu}|f(t)|\,dt, \qquad \int_1^\infty e^{-at}|f(t)|\,dt$$

be convergent, and let $f(t)$ *be of bounded variation in the neighbourhood of* $t = x$. *Let*

$$F(z) = \tfrac{1}{2}\int_0^\infty H_\nu^{(1)}(tz)(tz)^{\frac{1}{2}}f(t)\,dt, \tag{4.1}$$

where $\mathbf{I}(z) \geqslant a$. *Then*

$$\tfrac{1}{2}\{f(x+0)+f(x-0)\} = \lim_{\lambda\to\infty} \tfrac{1}{2}\int_{ia-\lambda}^{ia+\lambda} \{J_\nu(xz)-i\mathbf{H}_\nu(xz)\}(xz)^{\frac{1}{2}}F(z)\,dz. \tag{4.2}$$

We have

$$H_\nu^{(1)}(z) = \left\{\left(\frac{2}{\pi z}\right)^{\frac{1}{2}} + \frac{K}{z^{\frac{3}{2}}} + O\left(\frac{1}{|z|^{\frac{5}{2}}}\right)\right\} e^{i(z-\frac{1}{2}\nu\pi-\frac{1}{4}\pi)} \tag{4.3}$$

for z large, $-\pi < \arg z < 2\pi$. Hence $F(z)$ is regular for $\mathbf{I}(z) \geqslant a$, and continuous up to $\mathbf{I}(z) = a$. Also $F(z) \to 0$ as $z \to \infty$ uniformly in any strip $a \leqslant \alpha \leqslant \mathbf{I}(z) \leqslant \beta$, by the Riemann–Lebesgue theorem.

Let x be a given positive number, and let X be so large that $X \geqslant x+1$, and

$$\int_X^\infty e^{-at}|f(t)|\,dt \leqslant \epsilon.$$

Let $$F(z) = \tfrac{1}{2}\left(\int_0^X + \int_X^\infty\right) H_\nu^{(1)}(tz)(tz)^{\frac{1}{2}}f(t)\,dt = F_1(z)+F_2(z),$$

and

$$\begin{aligned}\phi(\lambda) &= \tfrac{1}{2}\int_{ia-\lambda}^{ia+\lambda} \{J_\nu(xz)-i\mathbf{H}_\nu(xz)\}(xz)^{\frac{1}{2}}\{F_1(z)+F_2(z)\}\,dz \\ &= \phi_1(\lambda)+\phi_2(\lambda) \quad \text{(say)}.\end{aligned}$$

Now

$$J_\nu(z)-i\mathbf{H}_\nu(z) = \left\{\left(\frac{2}{\pi z}\right)^{\frac{1}{2}} + \frac{K'}{z^{\frac{3}{2}}} + O\left(\frac{1}{|z|^{\frac{5}{2}}}\right)\right\} e^{-i(z-\frac{1}{2}\nu\pi-\frac{1}{4}\pi)} - ib_{0,\nu}z^{\nu-1} + O(|z|^{\nu-3}) \tag{4.4}$$

for $|z|$ large and $0 \leqslant \arg z \leqslant \pi$. If $0 \leqslant \arg z \leqslant \frac{1}{2}\pi$, this follows from (2.1) and the usual asymptotic formulae for J_ν and Y_ν. It is also true for $\frac{1}{2}\pi \leqslant \arg z \leqslant \pi$, as may be seen on putting

$$z = \zeta e^{i\pi}, \quad J_\nu(z) = e^{i\pi\nu} J_\nu(\zeta), \quad \mathbf{H}_\nu(z) = -e^{i\pi\nu}\mathbf{H}_\nu(\zeta),$$

and using the formula corresponding to (4.4) for $J_\nu(\zeta)+i\mathbf{H}_\nu(\zeta)$.

The leading terms in (4.3) and (4.4) contribute to $\phi_2(\lambda)$

$$\frac{1}{2\pi}\int_{ia-\lambda}^{ia+\lambda} e^{-ixz}\,dz \int_X^\infty e^{itz} f(t)\,dt = \frac{1}{\pi}\int_X^\infty \frac{e^{-a(t-x)}}{t-x}\sin\{(t-x)\lambda\} f(t)\,dt$$

$$= O\left\{\int_X^\infty e^{-at}|f(t)|\,dt\right\} = O(\epsilon),$$

uniformly with respect to λ. Similar arguments apply to the other terms, and it follows that $\phi_2(\lambda) = O(\epsilon)$, by choice of X, for all λ.

Now $F_1(z)$ is regular in the upper half-plane, and on the real axis except at $z = 0$. Consider the integral

$$\tfrac{1}{2}\int \{J_\nu(xz) - i\mathbf{H}_\nu(xz)\}(xz)^{\frac{1}{2}} F_1(z)\,dz$$

taken round the rectangle with vertices at $\pm\lambda$, $ia\pm\lambda$, with an indentation above the origin. The integrals along the vertical sides tend to 0 as $\lambda \to \infty$, and the integral round the indentation tends to 0 with the radius. Hence

$$\lim_{\lambda\to\infty} \phi_1(\lambda) = \tfrac{1}{2}\int_0^\infty \{J_\nu(xu) - i\mathbf{H}_\nu(xu)\}(xu)^{\frac{1}{2}} F_1(u)\,du +$$

$$+ \tfrac{1}{2}\int_0^\infty \{J_\nu(-xu) - i\mathbf{H}_\nu(-xu)\}(-xu)^{\frac{1}{2}} F_1(-u)\,du, \quad (4.5)$$

where, for instance, $J_\nu(-xu)$ denotes the value obtained from the real value of $J_\nu(xu)$ by continuous variation in the upper half-plane. Then

$$H_\nu^{(1)}(-u) = -e^{-\nu\pi i} H_\nu^{(2)}(u) = -e^{-\nu\pi i}\{J_\nu(u) - iY_\nu(u)\},$$

$$J_\nu(-u) = e^{\nu\pi i} J_\nu(u), \qquad \mathbf{H}_\nu(-u) = -e^{\nu\pi i}\mathbf{H}_\nu(u).$$

Hence the above right-hand side is equal to

$$\tfrac{1}{2}x^{\frac{1}{2}}\int_0^\infty J_\nu(xu)\{u^{\frac{1}{2}}F_1(u) + e^{\nu\pi i}(-u)^{\frac{1}{2}}F_1(-u)\}\,du -$$

$$-\tfrac{1}{2}ix^{\frac{1}{2}}\int_0^\infty \mathbf{H}_\nu(xu)\{u^{\frac{1}{2}}F_1(u) - e^{\nu\pi i}(-u)^{\frac{1}{2}}F_1(u)\}\,du,$$

where

$$u^{\frac{1}{2}}F_1(u)+e^{\nu\pi i}(-u)^{\frac{1}{2}}F_1(-u) = \tfrac{1}{2}u^{\frac{1}{2}}\int_0^X \{H_\nu^{(1)}(tu)+H_\nu^{(2)}(tu)\}(tu)^{\frac{1}{2}}f(t)\,dt$$

$$= u^{\frac{1}{2}}\int_0^X J_\nu(tu)(tu)^{\frac{1}{2}}f(t)\,dt,$$

and similarly

$$u^{\frac{1}{2}}F_1(u)-e^{\nu\pi i}(-u)^{\frac{1}{2}}F_1(-u) = iu^{\frac{1}{2}}\int_0^X Y_\nu(tu)(tu)^{\frac{1}{2}}f(t)\,dt.$$

Hence

$$\lim_{\lambda\to\infty}\phi_1(\lambda) = \tfrac{1}{2}\int_0^\infty J_\nu(xu)(xu)^{\frac{1}{2}}\,du\int_0^X J_\nu(tu)(tu)^{\frac{1}{2}}f(t)\,dt+$$

$$+\tfrac{1}{2}\int_0^\infty \mathbf{H}_\nu(xu)(xu)^{\frac{1}{2}}\,du\int_0^X Y_\nu(tu)(tu)^{\frac{1}{2}}f(t)\,dt$$

$$= \tfrac{1}{2}\{f(x+0)+f(x-0)\} \tag{4.6}$$

by (1.5) and (1.6). The theorem now follows.

5. The case $\nu \geqslant \frac{1}{2}$. We now take

$$F(z) = \frac{1}{2}\int_0^\infty \left\{H_\nu^{(1)}(tz)-i\sum_{0\leqslant m<\frac{1}{2}\nu-\frac{1}{4}}\frac{a_{m,\nu}}{(tz)^{\nu-2m}}\right\}(tz)^{\frac{1}{2}}f(t)\,dt, \tag{5.1}$$

where the sum consists of those terms of the principal part of $Y_\nu(tz)$ at the origin for which $\nu-2m > \frac{1}{2}$. The integral converges at the origin if $\int_0^1 |f(t)|\,dt$ is finite, and at infinity if

$$\int_1^\infty t^{2[\frac{1}{4}\nu-\frac{1}{4}]-\nu+\frac{1}{2}}|f(t)|\,dt$$

is convergent, except that, if $\frac{1}{2}\nu-\frac{1}{4}$ is an integer, the condition is that

$$\int_1^\infty t^{-2}|f(t)|\,dt$$

should be convergent. This formula cannot be applied to functions which are exponentially large at infinity, unless they satisfy special conditions at the origin so that the extra terms can be omitted.

The expansion formula is now

$$\tfrac{1}{2}\{f(x+0)+f(x-0)\}$$

$$= \frac{1}{2}\int_{ia-\infty}^{ia+\infty}\left\{J_\nu(xz)-i\mathbf{H}_\nu(xz)+i\sum_{0\leqslant m\leqslant\nu-\frac{1}{2}}\frac{b_{m,\nu}}{(xz)^{2m+1-\nu}}\right\}(xz)^{\frac{1}{2}}F(z)\,dz, \tag{5.2}$$

where the $b_{m,\nu}$ are those of (2.1). This can be derived from (4.2) by repeated integrations by parts. We have

$$\frac{d}{dz}\left[z^{\nu}\left\{H_{\nu}^{(1)}(tz)-i\sum_{m<\frac{1}{2}\nu-\frac{1}{4}}\frac{a_{m,\nu}}{(tz)^{\nu-2m}}\right\}\right]=tz^{\nu}\left\{H_{\nu-1}^{(1)}(tz)-i\sum_{m<\frac{1}{2}\nu-\frac{5}{4}}\frac{a_{m,\nu-1}}{(tz)^{\nu-1-2m}}\right\} \tag{5.3}$$

and

$$\int z^{-\nu+1}\left\{J_{\nu}(xz)-i\mathbf{H}_{\nu}(xz)+i\sum_{m\leqslant\nu-\frac{1}{2}}\frac{b_{m,\nu}}{(xz)^{2m+1-\nu}}\right\}dz$$
$$=-\frac{z^{-\nu+1}}{x}\left\{J_{\nu-1}(xz)-i\mathbf{H}_{\nu-1}(xz)+i\sum_{m\leqslant\nu-\frac{3}{2}}\frac{b_{m,\nu-1}}{(xz)^{2m+2-\nu}}\right\}. \tag{5.4}$$

Hence the right-hand side of (5.2) is

$$-\frac{1}{2x^{\frac{1}{2}}}\left[\left\{J_{\nu-1}(xz)-i\mathbf{H}_{\nu-1}(xz)+i\sum_{m\leqslant\nu-\frac{3}{2}}\frac{b_{m,\nu-1}}{(xz)^{2m+2-\nu}}\right\}z^{\frac{1}{2}}F(z)\right]_{ia-\infty}^{ia+\infty}+$$
$$+\frac{1}{2x^{\frac{1}{2}}}\int_{ia-\infty}^{ia+\infty}\left\{J_{\nu-1}(xz)-i\mathbf{H}_{\nu-1}(xz)+i\sum_{m\leqslant\nu-\frac{3}{2}}\frac{b_{m,\nu-1}}{(xz)^{2m+2-\nu}}\right\}F_1(z)\,dz,$$

where

$$F_1(z)=\frac{1}{2}\int_0^{\infty}\left\{H_{\nu-1}^{(1)}(tz)-i\sum_{m<\frac{1}{2}\nu-\frac{5}{4}}\frac{a_{m,\nu-1}}{(tz)^{\nu-2m-1}}\right\}(tz)^{\frac{1}{2}}tf(t)\,dt.$$

The integrated terms vanish if $F(ia\pm\lambda)\to 0$, and it is easily verified that this is so under the above conditions. In the case $\frac{1}{2}\leqslant\nu<\frac{3}{2}$, the result now follows by (4.2) with $f(t)$ replaced by $tf(t)$. If $\nu\geqslant\frac{3}{2}$, further integrations by parts are necessary. We obtain a second integrated term

$$-\frac{z^{\frac{1}{2}}}{2x^{\frac{3}{2}}}\left\{J_{\nu-2}(xz)-i\mathbf{H}_{\nu-2}(xz)+i\sum_{m<\nu-\frac{5}{2}}\frac{b_{m,\nu-2}}{(xz)^{2m+3-\nu}}\right\}F_1(z).$$

This tends to 0 if $F_1(ia\pm\lambda)\to 0$, and this is easily verified; in fact the situation becomes easier at each integration by parts.

The formulae can be varied to a certain extent according to the conditions which the function $f(t)$ satisfies. Suppose, for example, that $\frac{1}{2}\leqslant\nu<\frac{3}{2}$, and that

$$\int_0^1 t^{\frac{1}{2}-\nu}|f(t)|\,dt$$

is convergent. Then the term with $m=0$ which occurs in (5.1) can be omitted, and the condition at infinity can then be replaced by

that of § 4. If both sets of conditions are satisfied, both formulae are of course valid. In fact they differ by a multiple of

$$\int_{ia-\infty}^{ia+\infty} \left\{J_\nu(xz) - i\mathbf{H}_\nu(xz) + \frac{ib_{0,\nu}}{(xz)^{1-\nu}}\right\} z^{1-\nu}\,dz.$$

The integrand is regular at $z = 0$, and on moving the line of integration downwards it is seen that the integral vanishes.

6. Next suppose that $f(t)$ is defined over $(-\infty, \infty)$. In the case $-\frac{1}{2} \leqslant \nu < \frac{1}{2}$, let

$$F_+(z) = \tfrac{1}{2}\int_0^\infty H_\nu^{(1)}(tz)(tz)^{\frac{1}{2}} f(t)\,dt \quad (\mathbf{I}(z) > 0), \tag{6.1}$$

$$F_-(z) = \tfrac{1}{2}\int_{-\infty}^0 H_\nu^{(1)}(tz)(tz)^{\frac{1}{2}} f(t)\,dt \quad (\mathbf{I}(z) < 0). \tag{6.2}$$

In the second integral, $\arg tz$ varies from 0 to π, $H_\nu^{(1)}(tz)$ taking the value which is exponentially small for large negative z. The representation formula is then

$$\begin{aligned}\tfrac{1}{2}\{f(x+0)+f(x-0)\} = \tfrac{1}{2}\int_{ia-\infty}^{ia+\infty} \{J_\nu(xz) - i\mathbf{H}_\nu(xz)\}(xz)^{\frac{1}{2}} F_+(z)\,dz + \\ + \tfrac{1}{2}\int_{ib-\infty}^{ib+\infty} \{J_\nu(xz) - i\mathbf{H}_\nu(xz)\}(xz)^{\frac{1}{2}} F_-(z)\,dz,\end{aligned} \tag{6.3}$$

where $b < 0 < a$. If $x > 0$, the first term on the right converges to $\frac{1}{2}\{f(x+0)+f(x-0)\}$ as before. The second term converges to zero. The discussion of this is equivalent to the discussion of the first term with $x < 0$. This is equivalent to discussing (4.5) with i replaced by $-i$. This means that the two terms in (4.6) are subtracted instead of added. The result is therefore zero. Similarly (6.3) holds for $x < 0$. A similar extension may be made of the formulae with $\nu \geqslant \frac{1}{2}$.

7. Functions of order $n+\frac{1}{2}$

If $\nu = n+\frac{1}{2}$, the O-term in (2.1), with $p = n+1$, vanishes identically. We can therefore write the result as

$$\begin{aligned}\tfrac{1}{2}\{f(x+0)+f(x-0)\} = \tfrac{1}{2}\int_{ia-\infty}^{ia+\infty} H_{n+\frac{1}{2}}^{(2)}(xz)(xz)^{\frac{1}{2}} F_+(z)\,dz + \\ + \tfrac{1}{2}\int_{ib-\infty}^{ib+\infty} H_{n+\frac{1}{2}}^{(2)}(xz)(xz)^{\frac{1}{2}} F_-(z)\,dz,\end{aligned} \tag{7.1}$$

where $F_+(z)$ is (5.1), with $\nu = n+\frac{1}{2}$, and $F_-(z)$ is the corresponding integral over $(-\infty, 0)$. The conditions for this are

$$\int_0^1 |f(t)|\,dt + \int_1^\infty t^{-2}|f(t)|\,dt < \infty \quad (n \text{ even}),$$

$$\int_0^1 t|f(t)|\,dt + \int_1^\infty t^{-1}|f(t)|\,dt < \infty \quad (n \text{ odd}).$$

8. Uniqueness theorem. We shall consider finally the question of the uniqueness of the representation of a function in the form (6.3), or the corresponding form with $\nu \geqslant \frac{1}{2}$. The result in the case $\nu = -\frac{1}{2}$ is given by Theorem 141 of my *Fourier Integrals*. In the case $-\frac{1}{2} < \nu < \frac{1}{2}$, we use the formula

$$\int_0^\infty H_\nu^{(1)}(x\zeta)\{J_\nu(xz) - i\mathbf{H}_\nu(xz)\}x\,dx = \frac{2z^\nu}{\pi i\zeta^{\nu+1}(z-\zeta)}, \tag{8.1}$$

where $\mathbf{I}(z) < \mathbf{I}(\zeta)$. It is easily proved that

$$\int_0^\infty K_\nu(ax)\{J_\nu(xz) - i\mathbf{H}_\nu(xz)\}x\,dx = \frac{z^\nu}{a^{\nu+1}(a+iz)}, \tag{8.2}$$

by substituting the power series for J_ν and $\mathbf{H}_\nu$, integrating term by term, and using 13.21(8) of Watson's *Theory of Bessel Functions*. This is justified by absolute convergence if a is real and positive and $|z| < a$. Since

$$H_\nu^{(1)}(iax) = (2/\pi i)e^{-\frac{1}{2}\nu\pi i}K_\nu(ax),$$

(8.1) is equivalent to (8.2) with $\zeta = ia$. This proves (8.1) if $\arg\zeta = \frac{1}{2}\pi$, $|z| < |\zeta|$. The general result then follows by analytic continuation.

THEOREM. *Let* $-\frac{1}{2} \leqslant \nu < \frac{1}{2}$, *and let*

$$\int_{ia-\infty}^{ia+\infty} \{J_\nu(xz) - i\mathbf{H}_\nu(xz)\}(xz)^{\frac{1}{2}}\phi(z)\,dz + \int_{ib-\infty}^{ib+\infty} \{J_\nu(xz) - i\mathbf{H}_\nu(xz)\}(xz)^{\frac{1}{2}}\psi(z)\,dz = 0 \tag{8.3}$$

for $a_1 < a < a_2$, $b_1 < b < b_2$, *and all real* x, *where* $b_2 < 0 < a_1$. *Let* $\phi(z)$ *and* $\psi(z)$ *be analytic functions regular in these strips, tending to* 0 *at infinity in them, and absolutely integrable over their respective ranges.*

Then $z^{\nu+\frac{1}{2}}\phi(z)$ *and* $z^{\nu+\frac{1}{2}}\psi(z)$ *are regular for* $b_1 < \mathbf{I}(z) < a_2$, *their sum is* 0 *in this strip, and* $\phi(z) \to 0$ *as* $\mathbf{R}(z) \to \infty$, *uniformly in any interior strip.*

These conditions are obviously sufficient for (8.3) to vanish. To prove that they are necessary, multiply (8.3) by $x^{\frac{1}{2}}H_\nu^{(1)}(x\zeta)$, where $a < \mathbf{I}(\zeta) < a_2$, and integrate with respect to x over $(0,\infty)$. By (8.1) we obtain

$$\int_{ia-\infty}^{ia+\infty} \frac{z^{\nu+\frac{1}{2}}}{z-\zeta}\phi(z)\,dz + \int_{ib-\infty}^{ib+\infty} \frac{z^{\nu+\frac{1}{2}}}{z-\zeta}\psi(z)\,dz = 0.$$

The inversion is justified by absolute convergence. The theorem of residues then gives

$$2\pi i\zeta^{\nu+\frac{1}{2}}\phi(\zeta) + \int_{i\alpha-\infty}^{i\alpha+\infty} \frac{z^{\nu+\frac{1}{2}}}{z-\zeta}\phi(z)\,dz + \int_{i\beta-\infty}^{i\beta+\infty} \frac{z^{\nu+\frac{1}{2}}}{z-\zeta}\psi(z)\,dz = 0, \qquad (8.4)$$

where $\mathbf{I}(\zeta) < \alpha < a_2$, and β is any number in (b_1, b_2). It follows that $\zeta^{\nu+\frac{1}{2}}\phi(\zeta)$ is regular in $\beta < \mathbf{I}(\zeta) < \alpha$. Similarly by integrating over $(-\infty, 0)$ we obtain (8.4) with $\phi(\zeta)$ replaced by $-\psi(\zeta)$. Hence $\phi(\zeta)+\psi(\zeta) = 0$. Finally

$$z^{\nu+\frac{1}{2}}/(z-\zeta) = O(|\zeta|^{\nu+\frac{1}{2}})$$

uniformly with respect to z, if $|\mathbf{I}(z-\zeta)| \geqslant A > 0$, and the order result follows.

In the case $\frac{1}{2} \leqslant \nu < \frac{3}{2}$, (8.3) is replaced by

$$\int_{ia-\infty}^{ia+\infty} \left\{J_\nu(xz) - i\mathbf{H}_\nu(xz) + \frac{ib_{0,\nu}}{(xz)^{1-\nu}}\right\}(xz)^{\frac{1}{2}}\phi(z)\,dz + \ldots = 0. \qquad (8.5)$$

Integrating by parts and using (5.4), we obtain

$$\int_{ia-\infty}^{ia+\infty} \{J_{\nu-1}(xz) - i\mathbf{H}_{\nu-1}(xz)\}(xz)^{\frac{1}{2}}\left\{\phi'(z) + \frac{\nu-\frac{1}{2}}{z}\phi(z)\right\}dz + \ldots = 0.$$

Since $\phi'(z) \to 0$ at infinity, if $\phi(z) \to 0$, the previous case gives

$$\phi'(z) + \frac{\nu-\frac{1}{2}}{z}\phi(z) + \psi'(z) + \frac{\nu-\frac{1}{2}}{z}\psi(z) = 0,$$

$$\phi(z)+\psi(z) = cz^{\frac{1}{2}-\nu}. \qquad (8.6)$$

In the case $\nu = \frac{1}{2}$, we have $c = 0$ since ϕ and ψ both tend to 0 in the strip, but this is not obvious if $\nu > \frac{1}{2}$.

The previous case shows that

$$z^{\nu-\frac{1}{2}}\left\{\phi'(z) + \frac{\nu-\frac{1}{2}}{z}\phi(z)\right\}$$

is regular at the origin, and so, by integration, is $z^{\nu-\frac{1}{2}}\phi(z)$. As regards behaviour at infinity, the previous theorem shows that

$$\chi(z) = \phi'(z)+(\nu-\tfrac{1}{2})\phi(z)/z$$

tends to 0 uniformly in the whole strip. Writing

$$z^{\nu-\frac{1}{2}}\phi(z) = z_0^{\nu-\frac{1}{2}}\phi(z_0)+\int_{z_0}^{z}\zeta^{\nu-\frac{1}{2}}\chi(\zeta)\,d\zeta,$$

where $\mathbf{R}(z-z_0) = 0$, $\mathbf{I}(z_0) = a$, $a_1 < a < a_2$, it follows that $\phi(z)$ tends to 0 uniformly in the whole strip.

In the case $\frac{3}{2} \leqslant \nu < \frac{5}{2}$, (8.6) again holds (no term involving $z^{\frac{5}{2}-\nu}$ can occur). The given relation then reduces to

$$\int\left\{J_\nu(xz)-i\mathbf{H}_\nu(xz)+\frac{ib_{0,\nu}}{(xz)^{1-\nu}}+\frac{ib_{1,\nu}}{(xz)^{3-\nu}}\right\}(xz)^{\frac{1}{2}}\phi(z)\,dz-$$

$$-Cx^{\frac{1}{2}}\int_{ib-\infty}^{ib+\infty}\left\{J_\nu(xz)-i\mathbf{H}_\nu(xz)+\frac{ib_{0,\nu}}{(xz)^{1-\nu}}+\frac{ib_{1,\nu}}{(xz)^{3-\nu}}\right\}z^{1-\nu}\,dz = 0,$$

where the first integral is taken round a closed loop containing the origin. The second term is 0 for $x \neq 0$. If $x > 0$, this follows on moving the line of integration downwards, and, if $x < 0$, on moving it upwards and noting that the residue at $z = 0$ is 0. It follows that the other term must also vanish. Now the previous argument shows that

$$z^{\nu-\frac{3}{2}}\left\{\phi'(z)+\frac{\nu-\frac{1}{2}}{z}\phi(z)\right\} = \omega(z)$$

is regular at $z = 0$. Hence

$$z^{\nu-\frac{1}{2}}\phi(z) = A+\tfrac{1}{2}z^2\omega(0)+\ldots,$$

i.e.

$$\phi(z) = Az^{\frac{1}{2}-\nu}+Bz^{\frac{5}{2}-\nu}+\ldots,$$

in which there is no term $z^{\frac{3}{2}-\nu}$. The general result is now clear. If $n+\frac{1}{2} \leqslant \nu < n+\frac{3}{2}$, the functions ϕ and ψ are related by

$$\phi(z)+\psi(z) = C_0 z^{\frac{1}{2}-\nu}+\ldots+C_p z^{2p+\frac{1}{2}-\nu},$$

where $2p+\frac{1}{2} < \nu$; and, in the neighbourhood of the origin,

$$\phi(z) = z^{\frac{1}{2}-\nu}(A+Bz^2+\ldots+Kz^{2n}+Lz^{2n+1}+\ldots),$$

and $\psi(z)$ is of the same form.

A NOTE ON SOME ELEMENTARY TAUBERIAN THEOREMS

By H. R. PITT (*Belfast*)

[Received 8 December 1947]

1. Introduction

THE classical Tauberian theorems can all be expressed in the form that, if $k(x)$ is integrable in $(-\infty, \infty)$ and

$$g(u) = \int_{-\infty}^{\infty} k(u-y)s(y)\,dy, \tag{1.1}$$

the convergence of $g(x)$ as $x \to \infty$ implies that of $s(x)$ whenever $s(x)$ is bounded and satisfies a suitable *Tauberian condition*. A sufficient condition, whatever the form of $k(x)$, is that $s(y)$ be differentiable and $s'(y) = o(1)$ as $y \to \infty$, and this condition, although differentiability is an unnecessarily stringent requirement, is of the correct order in the sense that the condition $s'(y) = o[p(y)]$ for an unbounded $p(y)$ is not sufficient.

The Tauberian theorem with the condition $s'(y) = o(1)$ may be called an *elementary* theorem to distinguish it from the much deeper theorem with the Tauberian condition $s'(y) = O(1)$, which holds when $k(x)$ satisfies certain extra conditions. A similar distinction can usually be made even when the Tauberian condition is not expressed in terms of the derivative of $s(y)$.

We are concerned in this note with elementary Tauberian theorems for the transformation

$$g(u) = \int_{-\infty}^{\infty} k(u,y)s(y)\,dy, \tag{1.2}$$

where, except in Theorem 3, the kernel $k(u,y)$ satisfies the conditions:

$$k(u,y) \geqslant 0, \tag{1.3}$$

$$\int k(u,y)\,dy = 1 \quad \text{for all real } u, \tag{1.4}$$

$$\lim_{u\to\infty} \int_{-\infty}^{Y} k(u,y)\,dy = 0 \quad \text{for every fixed } Y. \tag{1.5}$$

[Here, and in what follows, I omit limits of integration when they are $-\infty$ and ∞.]

There is no difficulty in proving that such a transformation is consistent in the sense that the existence of $\lim_{y\to\infty} s(y)$ implies the existence of $\lim_{y\to\infty} g(y)$ with the same value.

2. Tauberian conditions

It is convenient to recapitulate in the form required in the present problem the analysis of Tauberian conditions which was given in a previous paper† in connexion with transformations of type (1.1).

I say that a real or complex function $s(x)$ satisfies condition $T[\delta, \theta, \xi]$ if functions $\delta(\epsilon, x)$, $\theta(\epsilon, x)$, $\xi(\epsilon, x)$ can be defined for all x and $\epsilon \geqslant 0$ so that

$$\delta(\epsilon, x) > 0, \tag{2.1}$$

$$|\xi - x| \leqslant \delta, \tag{2.2}$$

and

$$R\{e^{i\theta} s(y)\} \geqslant |s(x)| - \epsilon \tag{2.3}$$

for

$$|\xi - y| < \delta.$$

I say that $s(x)$ satisfies the Tauberian condition $T_\alpha[\delta, \xi]$ if $\delta(\epsilon, x)$, $\xi(\epsilon, x, \phi)$ can be defined when x is real, $\epsilon \geqslant 0$, $-\pi \leqslant \phi < \pi$, so that (2.1), (2.2) are satisfied and also

$$R\{e^{i\phi}[s(y) - s(x)]\} \geqslant -\epsilon \quad (|\xi - y| < \delta). \tag{2.4}$$

I say that $s(x)$ satisfies the Tauberian condition $T_\beta[\delta, \xi]$ if $\xi = \xi(\epsilon, x)$ is independent of ϕ. Condition (2.4) is then equivalent to

$$|s(y) - s(x)| \leqslant \epsilon \quad (|\xi - y| < \delta). \tag{2.5}$$

Conditions T_α, T_β correspond respectively to the one-sided and two-sided conditions in the classical theorems. A fuller discussion is in the paper to which I refer above.

3. Elementary Tauberian theorems

THEOREM 1. *Suppose that $k(u, y)$ satisfies conditions* (1.3), (1.4), *and* (1.5), *that $s(x)$ is bounded, and that*

$$g(u) = \int k(u, y) s(y)\, dy. \tag{3.1}$$

† H. R. Pitt, 'General Tauberian theorems', *Proc. London Math. Soc.* (2), 44 (1938), 243–88.

For simplicity, we suppose that the function $\mu(\epsilon, x)$ of the earlier paper vanishes.

Suppose that $s(x)$ satisfies condition $T(\delta, \theta, \xi)$, that $\epsilon > 0$, and that

$$\Lambda(\epsilon) = \lim_{v\to\infty} \overline{\lim_{x\to\infty}} \text{ u.b.}_{u>v} \int_{\xi-\delta}^{\xi+\delta} k(u,y)\,dy > \tfrac{1}{2}. \tag{3.2}$$

Then, if

$$S = \overline{\lim_{x\to\infty}} |s(x)|, \qquad G = \overline{\lim_{x\to\infty}} |g(x)|,$$

we have

$$S[2\Lambda(\epsilon)-1] \leqslant G+\epsilon \tag{3.3}$$

and, if $\Lambda(+0) > \frac{1}{2}$,

$$S \leqslant \frac{G}{2\Lambda(+0)-1}. \tag{3.4}$$

If we write (3.1) as

$$g(u) = \int_{\xi-\delta}^{\xi+\delta} k(u,y)s(y)\,dy + \left\{\int_{-\infty}^{\xi-\delta} + \int_{\xi+\delta}^{\infty}\right\} k(u,y)s(y)\,dy, \tag{3.5}$$

and use (2.3), we get

$$\begin{aligned} |g(u)| &\geqslant R\{g(u)e^{i\theta}\} \\ &\geqslant (|s(x)|-\epsilon)\int_{\xi-\delta}^{\xi+\delta} k(u,y)\,dy - \left\{\int_{-\infty}^{\xi-\delta} + \int_{\xi+\delta}^{\infty}\right\} k(u,y)|s(y)|\,dy. \end{aligned} \tag{3.6}$$

Writing

$$S(Y) = \text{u.b.}_{y<Y} |s(y)|, \qquad S' = S(-\infty) = \text{u.b.}_{-\infty<y<\infty} s|(y)|,$$

we see easily from (1.3), (1.4), and (3.6) that, for any v and Y, we have

$$\begin{aligned} &\text{u.b.}_{u<v} |g(u)| \\ &\geqslant [S+S(Y)] \lim_{x\to\infty} \text{u.b.}_{u>v} \int_{\xi-\delta}^{\xi+\delta} k(u,y)\,dy - S(Y) - S' \text{ u.b.}_{u>v} \int_{-\infty}^{Y} k(u,y)\,dy - \epsilon. \end{aligned}$$

Letting $v \to \infty$, we get

$$G \geqslant [S+S(Y)]\Lambda(\epsilon) - S(Y) - \epsilon,$$

and the conclusion follows as we let $Y \to \infty$.

THEOREM 2. *Suppose that the conditions of Theorem* 1 *are satisfied, that $s(x)$ satisfies the stronger condition T_α, and that*

$$\Lambda(\epsilon) = 1 \tag{3.7}$$

for every positive ϵ.

Then the convex cover† of the set of limit points of $s(x)$ coincides with the convex cover of the limit points of $g(x)$.

† The product of all closed convex sets containing the limit points of $s(x)$.

If we define S' as in Theorem 1, and use (2.4), we have

$$R\{e^{i\phi}[g(u)-s(x)]\} = R\left[\left\{\int_{-\infty}^{\xi-\delta} + \int_{\xi+\delta}^{\infty}\right\}k(u,y)e^{i\phi}[s(y)-s(x)]\,dy\right] +$$

$$+R\left\{\int_{\xi-\delta}^{\xi+\delta} k(u,y)e^{i\phi}[s(y)-s(x)]\,dy\right\} \geqslant 2S'\left[\int_{\xi-\delta}^{\xi+\delta} k(u,y)\,dy-1\right]-\epsilon.$$

It then follows from (3.7) that

$$\lim_{v\to\infty}\overline{\lim_{x\to\infty}}\ \text{u.b.}_{u>v}\ R\{e^{i\phi}[g(u)-s(x)]\} \geqslant -\epsilon,$$

and, since this is true for every positive ϵ, the left-hand side must vanish for every real ϕ. This shows that every limit point of $s(x)$ lies in the convex cover of the limit points of $g(x)$, and the fact that the limit points of $g(x)$ lie in the convex cover of limit points of $s(x)$ follows almost trivially from (1.3), (1.4), and (1.5).

If the condition T_α is replaced by T_β, we need no longer suppose that $k(u,y) \geqslant 0$, and we have a stronger conclusion.

THEOREM 3. *Suppose that $k(u,y)$ satisfies* (1.4) *and* (1.5), *that $s(x)$ is bounded and satisfies condition T_β, and that*

$$\int |k(u,y)|\,dy \leqslant M \quad \textit{for all } u, \tag{3.8}$$

$$\lim_{v\to\infty}\overline{\lim_{x\to\infty}}\ \text{l.b.}_{u>v}\left\{\int_{-\infty}^{\xi-\delta} + \int_{\xi+\delta}^{\infty}\right\}|k(u,y)|\,dy = 0. \tag{3.9}$$

Then every limit point of $s(x)$ is a limit point of $g(x)$.

Conditions (2.5) and (3.8) allow us to write

$$|g(u)-s(x)| \leqslant 2S'\left\{\int_{-\infty}^{\xi-\delta} + \int_{\xi+\delta}^{\infty}\right\}|k(u,y)|\,dy+\epsilon M,$$

and it follows from (3.9) that

$$\lim_{v\to\infty}\overline{\lim_{x\to\infty}}\ \text{l.b.}_{u>v}\ |g(u)-s(x)| \leqslant \epsilon,$$

which is enough to establish the conclusion.

ON SOME DIOPHANTINE EQUATIONS $y^2 = x^3+k$ WITH NO RATIONAL SOLUTIONS

By K. L. CHANG (*Nanchang*)

[Received 16 December 1947]

1. FUETER has found some values of k for which the Diophantine equation

$$y^2 = x^3+k$$

has no rational solutions. Other values have been found by Mordell. By using Mordell's method, I have found* some other results as stated in the following theorem.

THEOREM. *The equation $y^2 = x^3+k$ has no rational solutions if k is a square-free positive integer, and*

(1) $$k \equiv 5 \pmod 8, \qquad k \equiv -3 \pmod 9,$$
$$\text{i.e.} \quad k \equiv -3 \pmod{72};$$

(2) *the class number of the field $R(\sqrt{k})$ is not divisible by 3;*

(3) $$u \not\equiv 0 \pmod 3, \qquad u \not\equiv \pm 2 \pmod 9,$$

where u is the least positive value of q satisfying the Pellian equation $p^2-kq^2 = \pm 4$;

(4) *the class number h of the field $R\{\sqrt{(-\frac{1}{3}k)}\}$ is not divisible by 3.*

(5) *α, β and γ, δ the respective integer solutions of the equations*

$$\tfrac{1}{4}(\alpha^2+\tfrac{1}{3}k\beta^2) = 2^h, \qquad \tfrac{1}{4}(\gamma^2+\tfrac{1}{3}k\delta^2) = 3^{2h}$$

satisfy the conditions:

(*a*) (i) $\alpha \not\equiv 0 \pmod 9$ *when h is odd,*

(ii) $\beta \not\equiv 0 \pmod 9$ *when h is even*;

(*b*) *when $h = 3n+1$,*

$$\{\alpha \pm \beta(\tfrac{1}{3}k)^2\}\delta \not\equiv \pm 2 \pmod 9,$$

and

$$\delta \not\equiv \pm 2 \pmod 9;$$

when $h = 3n-1$, $\{\alpha(\frac{1}{3}k) \pm \beta\}\delta \not\equiv \pm 2 \pmod 9$

and

$$\delta \not\equiv \pm 4(\tfrac{1}{3}k)^2 \pmod 9;$$

where all the signs are independent of each other.

* I am greatly obliged to Professor Mordell for suggesting this work to me and for his valuable comments.

A particular instance is $k = 357$, where the class number of the field $R(\sqrt{357})$ in (2) is 4, the value of u in (3) is 1, the value of h in (4) is 10, and the values of α, β, δ in (5) are 55, 3, 9720 respectively.

2. Let
$$y^2 = x^3+k \tag{I}$$
be the given equation.

Put
$$y = Y/Z',$$
where Y, Z' are integers and $(Y, Z') = 1$. Then
$$Z'(Y^2-Z'^2k) = Z'^3x^3.$$
Since $(Z', Y^2-Z'^2k) = 1$, each of the factors Z' and $Y^2-Z'^2k$ must be a perfect cube. Therefore we may put $Z' = Z^3$, and then
$$Y^2-kZ^6 = (Z^2x)^3.$$
Since Z^2x is an integer, we may put $Z^2x = X$. Hence
$$Y^2-kZ^6 = X^3, \tag{II}$$
where X, Y, Z are integers.

The original problem of solving (I) in rational numbers is changed into the problem of solving (II) in integers.

In (II), $(Y, Z) = 1$, since $(Y, Z') = 1$ and so $(Z, X) = 1$. Also $(Y, k) = 1$, for, if there exists a prime divisor p of (Y, k), then $p \mid Z$, since k is square-free. This is a contradiction.

Since $(Y, Z) = 1$, Y and Z cannot be both even. Since $k \equiv 5 \pmod 8$, Y and Z cannot be both odd. Hence X must be odd. From (II) we have
$$(Y+Z^3\sqrt{k})(Y-Z^3\sqrt{k}) = X_1^3.$$

In the quadratic field $R(\sqrt{k})$ the ideals $[Y+Z^3\sqrt{k}]$, $[Y-Z^3\sqrt{k}]$ are prime to one another since X is odd and Y is prime to kZ. Hence
$$[Y+Z^3\sqrt{k}] = X_1^3,$$
where X_1 must be a principal ideal in the field $R(\sqrt{k})$ from (2), and so we have
$$Y+Z^3\sqrt{k} = \epsilon(\tfrac{1}{2}A+\tfrac{1}{2}B\sqrt{k})^3,$$
where ϵ is a unit in the field $R(\sqrt{k})$ and $\frac{1}{2}A+\frac{1}{2}B\sqrt{k}$ is an integer in the same field: that is, A and B are integers either both odd or both even.

Also
$$\epsilon = \pm(\tfrac{1}{2}T+\tfrac{1}{2}u\sqrt{k})^n \quad (n = 0, \pm1, \pm2, \ldots),$$

where T and u are integers either both odd or both even. After absorbing cube factors in $(\frac{1}{2}A+\frac{1}{2}B\sqrt{k})^3$, we have either

$$\left.\begin{aligned} Y+Z^3\sqrt{k} &= (\tfrac{1}{2}T\pm\tfrac{1}{2}u\sqrt{k})(\tfrac{1}{2}A+\tfrac{1}{2}B\sqrt{k})^3 \\ X &= \tfrac{1}{4}(A^2-kB^2) \end{aligned}\right\}, \qquad \text{(III)}$$

or

$$\left.\begin{aligned} Y+Z^3\sqrt{k} &= (\tfrac{1}{2}A+\tfrac{1}{2}B\sqrt{k})^3 \\ X &= \tfrac{1}{4}(A^2-kB^2) \end{aligned}\right\}, \qquad \text{(IV)}$$

where A and B are integers, and $(A, B) = 1$ when A and B are both odd; $(\frac{1}{2}A, \frac{1}{2}B) = 1$ when A and B are both even, since $(Y, Z) = 1$.

3. Consider the first alternative (III). Then

$$16Z^3 = \pm u(A^3+3AB^2k)+T(3A^2B+B^3k).$$

We may take the positive sign only, on writing $-A$ for A, if need be. Taking remainders to modulus 9 we have

$$-2Z^3 \equiv uA^3+T(3A^2B+B^3k) \pmod 9.$$

I show that the above congruences are impossible when $u \not\equiv 0 \pmod 3$ and $u \not\equiv \pm 2 \pmod 9$.

First, $Z \not\equiv 0 \pmod 3$. For, if $Z \equiv 0 \pmod 3$, we get $A \equiv 0 \pmod 3$ and $B \equiv 0 \pmod 3$ since $T \not\equiv 0 \pmod 3$, but $(A, B) = 1$. Hence $Z \equiv \pm 1 \pmod 3$ and

$$-2 \equiv uA^3+T(3A^2B+B^3k) \pmod 9.$$

Clearly $A \not\equiv 0 \pmod 3$, for otherwise $-2 \equiv TB^3k \pmod 9$; also $B \not\equiv 0 \pmod 3$, for otherwise $-2 \equiv uA^3 \pmod 9$, and $u \not\equiv \pm 2 \pmod 9$. Hence $A^2 \equiv B^2 \pmod 3$, and so

$$\begin{aligned} 3A^2B+B^3k &\equiv 3B(A^2-B^2) \pmod 9 \\ &\equiv 0 \pmod 9, \end{aligned}$$

and then $-2 \equiv uA^3 \pmod 9$, which is impossible.

Therefore, when $u \not\equiv 0 \pmod 3$ and $u \not\equiv \pm 2 \pmod 9$, no integer solutions of (III) can be found.

4. Next we consider the second alternative (IV).

(4 A) *When A and B are both odd.* We then have

$$Y+Z^3\sqrt{k} = (\tfrac{1}{2}A+\tfrac{1}{2}B\sqrt{k})^3,$$

where $(A, B) = 1$.

Hence $$8Z^3 = 3B(A^2+\tfrac{1}{3}kB^2).$$

Since B is odd and $(B, A^2+\frac{1}{3}kB^2) = 1$, we have either

$$B = 9C^3, \quad \tfrac{1}{3}(A^2+\tfrac{1}{3}kB^2) = D^3,$$

or $$B = C^3, \quad \tfrac{1}{3}(A^2+\tfrac{1}{3}kB^2) = 9D^3 \quad \{B \not\equiv 0 \pmod 3\}.$$

(4 Aa) From the first case we have

$$\{\tfrac{1}{2}A+\tfrac{1}{2}B\sqrt{(-\tfrac{1}{3}k)}\}\{\tfrac{1}{2}A-\tfrac{1}{2}B\sqrt{(-\tfrac{1}{3}k)}\} = 2D^3.$$

Since $\{\frac{1}{2}A+\frac{1}{2}B\sqrt{(-\frac{1}{3}k)}, \frac{1}{2}A-\frac{1}{2}B\sqrt{(-\frac{1}{3}k)}\} = 1$ and $(D, \frac{1}{3}k) = 1$, we have

$$[\tfrac{1}{2}A+\tfrac{1}{2}B\sqrt{(-\tfrac{1}{3}k)}] = I_1 D_1^3, \qquad [\tfrac{1}{2}A-\tfrac{1}{2}B\sqrt{(-\tfrac{1}{3}k)}] = I_2 D_2^3,$$

where I_1, I_2, D_1, D_2 are ideals in the field $R\{\sqrt{(-\frac{1}{3}k)}\}$, $I_1 I_2 = [2]$, $D_1 D_2 = [D]$.

Let the class number h of the field $R\{\sqrt{(-\frac{1}{3}k)}\}$ be not divisible by 3. Then I_1^h, I_2^h, D_1^h, D_2^h are principal ideals, and, since the fundamental units are ± 1, we have

$$\{\tfrac{1}{2}A\pm\tfrac{1}{2}B\sqrt{(-\tfrac{1}{3}k)}\}^h = \{\tfrac{1}{2}\alpha\pm\tfrac{1}{2}\beta\sqrt{(-\tfrac{1}{3}k)}\}\{\tfrac{1}{2}E\pm\tfrac{1}{2}F\sqrt{(-\tfrac{1}{3}k)}\}^3.$$

Here α, β and E, F are pairs of integers and the integers in each pair are either both odd or both even, and also α and β satisfy

$$\tfrac{1}{4}(\alpha^2+\tfrac{1}{3}k\beta^2) = 2^h.$$

Since $B = 9C^3 \equiv 0 \pmod 9$, we have

$$\{\tfrac{1}{2}A+\tfrac{1}{2}B\sqrt{(-\tfrac{1}{3}k)}\}^h-\{\tfrac{1}{2}A-\tfrac{1}{2}B\sqrt{(-\tfrac{1}{3}k)}\}^h \equiv 0 \pmod 9,$$

and then

$$\{\tfrac{1}{2}\alpha\pm\tfrac{1}{2}\beta\sqrt{(-\tfrac{1}{3}k)}\}\{\tfrac{1}{2}E+\tfrac{1}{2}F\sqrt{(-\tfrac{1}{3}k)}\}^3- \\ -\{\tfrac{1}{2}\alpha\mp\tfrac{1}{2}\beta\sqrt{(-\tfrac{1}{3}k)}\}\{\tfrac{1}{2}E-\tfrac{1}{2}F\sqrt{(-\tfrac{1}{3}k)}\}^3 \equiv 0 \pmod 9.$$

Since $-\frac{1}{3}k \equiv 1 \pmod 3$, then $\sqrt{(-\frac{1}{3}k)} \equiv t \pmod 9$ is soluble for t, and we have

$$(\tfrac{1}{2}\alpha\pm\tfrac{1}{2}t\beta)(\tfrac{1}{2}F+\tfrac{1}{2}tF)^3-(\tfrac{1}{2}\alpha\mp\tfrac{1}{2}t\beta)(\tfrac{1}{2}E-\tfrac{1}{2}tF)^3 \equiv 0 \pmod 9.$$

We may exclude the cases $\frac{1}{2}F\pm\frac{1}{2}tF \equiv 0 \pmod 3$, $\frac{1}{2}\alpha\pm\frac{1}{2}t\beta \equiv 0 \pmod 3$, since then $A \equiv 0 \pmod 3$, but $(A, B) = 1$.

Hence we have

$$\frac{\alpha+t\beta}{\alpha-t\beta} \equiv \frac{(E\pm tF)^3}{(E\mp tF)^3} \equiv \pm 1 \pmod 9.$$

Then either $\alpha \equiv 0 \pmod 9$ or $\beta \equiv 0 \pmod 9$.

Taking remainders to modulus 9 on the equation

$$\tfrac{1}{4}(\alpha^2+\tfrac{1}{3}k\beta^2) = 2^h,$$

we obtain $$\tfrac{1}{4}(\alpha^2-t^2\beta^2) \equiv 2^h \pmod 9.$$

If $\alpha \equiv 0 \pmod 9$, we have

$$\tfrac{1}{4}t^2\beta^2 \equiv -2^h \pmod 9,$$

which is soluble only when h is odd. If $\beta \equiv 0 \pmod 9$, we have

$$\tfrac{1}{4}\alpha^2 \equiv 2^h \pmod 9$$

which is soluble only when h is even.

Therefore there are no solutions in this case if (i) $\alpha \not\equiv 0 \pmod 9$ when h is odd; (ii) $\beta \not\equiv 0 \pmod 9$ when h is even.

(4 Ab) From the second case, we have

$$[\{\tfrac{1}{2}A+\tfrac{1}{2}B\sqrt{(-\tfrac{1}{3}k)}\}\{\tfrac{1}{2}A-\tfrac{1}{2}B\sqrt{(-\tfrac{1}{3}k)}\}] = [2.3^2.D^3] = I_1 I_2 J_1^2 J_2^2 D_1^3 D_2^3,$$

where $I_1, I_2, J_1, J_2, D_1, D_2$ are ideals in the field $R\sqrt{(-\frac{1}{3}k)}$, $I_1 I_2 = [2]$, $J_1 J_2 = [3]$, and $D_1 D_2 = [D]$.

Let h be defined as before. The fundamental units are ± 1, and we have, with arbitrary choice of the $\pm$signs,

$$\{\tfrac{1}{2}A \pm \tfrac{1}{2}B\sqrt{(-\tfrac{1}{3}k)}\}^h = \{\tfrac{1}{2}\alpha \pm \tfrac{1}{2}\beta\sqrt{(-\tfrac{1}{3}k)}\}\{\tfrac{1}{2}\gamma \pm \tfrac{1}{2}\delta\sqrt{(-\tfrac{1}{3}k)}\}\{\tfrac{1}{2}E \pm \tfrac{1}{2}F\sqrt{(-\tfrac{1}{3}k)}\}^3,$$

where α, β and γ, δ and E, F are pairs of integers, and the integers in each pair are either both odd or both even; and also α, β and γ, δ respectively satisfy the two equations

$$\tfrac{1}{4}(\alpha^2+\tfrac{1}{3}k\beta^2) = 2^h, \qquad \tfrac{1}{4}(\gamma^2+\tfrac{1}{3}k\delta^2) = 3^{2h}.$$

Since $-\frac{1}{3}k \equiv t^2 \pmod{3^{2h}}$ is soluble in t, we have

$$\gamma^2 - t^2\delta^2 \equiv 0 \pmod{3^{2h}}.$$

We may exclude the cases $\alpha \equiv \beta \equiv 0 \pmod 3$, $\gamma \equiv \delta \equiv 0 \pmod 3$, and $E \equiv F \equiv 0 \pmod 3$ since in any one of these cases

$$\{\tfrac{1}{2}A+\tfrac{1}{2}B\sqrt{(-\tfrac{1}{3}k)}\}^h \equiv 0 \pmod 3,$$
$$\{\tfrac{1}{2}A-\tfrac{1}{2}B\sqrt{(-\tfrac{1}{3}k)}\}^h \equiv 0 \pmod 3$$

and then $B \equiv 0 \pmod 3$.

We may suppose that t is defined by

$$\gamma - t\delta \equiv 0 \pmod{3^{2h}}.$$

Then we have either

$$(\tfrac{1}{2}A-\tfrac{1}{2}Bt)^h \equiv 0 \pmod{3^{2h}}, \qquad \tfrac{1}{2}A-\tfrac{1}{2}Bt \equiv 0 \pmod 9;$$

or

$$(\tfrac{1}{2}A+\tfrac{1}{2}Bt)^h \equiv 0 \pmod{3^{2h}}, \qquad \tfrac{1}{2}A+\tfrac{1}{2}Bt \equiv 0 \pmod 9.$$

Hence

$$(\pm Bt)^h \equiv (\tfrac{1}{2}\alpha \pm \tfrac{1}{2}t\beta)\delta t(\tfrac{1}{2}E \pm \tfrac{1}{2}tF)^3 \pmod 9,$$

where all the $\pm$ signs may be chosen arbitrarily.

We may exclude the cases $\frac{1}{2}(\alpha \pm t\beta) \equiv 0 \pmod 3$, or $\frac{1}{2}E \pm \frac{1}{2}tF \equiv 0 \pmod 3$ since then $B \equiv 0 \pmod 3$, but we know that $B \not\equiv 0 \pmod 3$. Hence $(\frac{1}{2}E \pm \frac{1}{2}tF)^3 \equiv \pm 1 \pmod 9$ and, since $B = C^3 \equiv \pm 1 \pmod 9$,

$$\pm t^h \equiv (\tfrac{1}{2}\alpha \pm \tfrac{1}{2}t\beta)\delta t \pmod 9,$$

i.e.

$$\pm 2t^{h-1} \equiv (\alpha \pm t\beta)\delta \pmod 9.$$

If $h = 3n+1$,

$$\begin{aligned} \pm 2 &\equiv (\alpha \pm t\beta)\delta \pmod 9 \\ &\equiv (\alpha \pm \beta/t^2)\delta \pmod 9 \\ &\equiv \{\alpha \pm \beta(\tfrac{1}{3}k)^2\}\delta \pmod 9. \end{aligned}$$

If $h = 3n-1$,

$$\pm 2t^{-2} \equiv (\alpha \pm t\beta)\delta \pmod 9,$$

or

$$\begin{aligned} \pm 2 &\equiv (\alpha t^2 + \beta)\delta \pmod 9 \\ &\equiv \{\alpha(-\tfrac{1}{3}k) \pm \beta\}\delta \pmod 9, \end{aligned}$$

or

$$\pm 2 \equiv \{\alpha(\tfrac{1}{3}k) \pm \beta\}\delta \pmod 9.$$

Therefore there are no solutions in this case if

(i) $\{\alpha \pm \beta(\tfrac{1}{3}k)^2\}\delta \not\equiv \pm 2 \pmod 9$ when $h = 3n+1$;

(ii) $\{\alpha(\tfrac{1}{3}k) \pm \beta\}\delta \not\equiv \pm 2 \pmod 9$ when $h = 3n-1$;

where all the $\pm$ signs are independent of each other.

(4 B) *When A and B in* (IV) *are both even.* On writing $2A$ for A, $2B$ for B, the equations in (IV) become

$$\left.\begin{aligned} Y + Z^3\sqrt{k} &= (A + B\sqrt{k})^3 \\ X &= A^2 - kB^2 \end{aligned}\right\}, \tag{V}$$

where A and B are neither both odd nor both even, since X is odd; also $(A, B) = 1$, since $(Y, Z) = 1$. Hence, from (V),

$$Z^3 = 3B(A^2 + \tfrac{1}{3}kB^2).$$

Since $(B, A^2 + \tfrac{1}{3}kB^2) = 1$, we have either

$$B = 9C^3, \qquad A^2 + \tfrac{1}{3}kB^2 = D^3;$$

or

$$B = C^3, \qquad A^2 + \tfrac{1}{3}kB^2 = 9D^3.$$

(4 Ba) In the first case D is odd and $(D, \tfrac{1}{3}k) = 1$.

Hence

$$[\{A + \sqrt{(-\tfrac{1}{3}k)}B\}\{A - \sqrt{(-\tfrac{1}{3}k)}B\}] = D_1^3 D_2^3,$$

where D_1, D_2 are ideals in the field $R\{\sqrt{(-\tfrac{1}{3}k)}\}$, $D_1 D_2 = [D]$.

Let h be defined as before. Since the fundamental units are ± 1, we have

$$A \pm \sqrt{(-\tfrac{1}{3}k)}B = \{\tfrac{1}{2}E \pm \tfrac{1}{2}F\sqrt{(-\tfrac{1}{3}k)}\}^3,$$

where E, F are integers and are either both odd and also $(E, F) = 1$, or both even and also $(\tfrac{1}{2}E, \tfrac{1}{2}F) = 1$.

Suppose that E and F are both odd. Then

$$8B = F(3E^2 - \tfrac{1}{3}kF^2), \qquad B = 9C^3.$$

But $3E^2 - \tfrac{1}{3}kF^2 \equiv 4 \pmod 8$, so this case is impossible. Hence E and F are both even. On writing $2E_1$ for E, $2F_1$ for F we have

$$A \pm B\sqrt{(-\tfrac{1}{3}k)} = \{E_1 \pm F_1\sqrt{(-\tfrac{1}{3}k)}\}^3,$$

where E_1 and F_1 are integers and $(E_1, F_1) = 1$. Hence

$$B = 9C^3 = F_1(3E_1^2 - \tfrac{1}{3}kF_1^2).$$

There are three cases,

$F_1 = G^3 \not\equiv 0 \pmod 3$, $\quad 3E_1^2 - \tfrac{1}{3}kF_1^2 = 9H^3$;

$F_1 = 3G^3$, $\quad 3E_1^2 - \tfrac{1}{3}kF_1^2 = 3H^3$;

$F_1 = 9G^3$, $\quad 3E_1^2 - \tfrac{1}{3}kF_1^2 = H^3 \not\equiv 0 \pmod 3$.

We may rule out the first and the third cases, in which there are contradictions. From the second case we have

$$E_1^2 - kG^6 = H^3,$$

which is of the same form as the original equation

$$Y^2 - kZ^6 = X^3.$$

Since from (V)

$$|Y| = |A^3 + 3AB^2k| = |A| \,.\, |A^2 + 3B^2k| > |A| \\ = |E_1| \,.\, |E_1^2 - F_1^2k| \geqslant |E_1|,$$

$$|Z|^3 = |3B| \,.\, |A^2 + \tfrac{1}{3}kB^2| > |3B| = 3|F_1| \,.\, |3E_1^2 - \tfrac{1}{3}kF_1^2| \geqslant 3|F_1| \\ = 3|3G^3| > |2G|^3,$$

we have $|E_1| < |Y|$, $|G| < |\tfrac{1}{2}Z|$.

We continue this process and either reach an equation

$$E_n^2 - kG_n^6 = H_n^3$$

with $|G_n| < 1$, which is a contradiction; or come to the second alternative in (4 B), namely,

$$B = C^3, \qquad A^2 + \tfrac{1}{3}kB^2 = 9D^3,$$

where $(A, B) = 1$.

(4 Bb) In $B = C^3$, $A^2 + \tfrac{1}{3}kB^2 = 9D^3$, we know that D is odd and $(D, \tfrac{1}{3}k) = 1$. Hence

$$[\{A + B\sqrt{(-\tfrac{1}{3}k)}\}\{A - B\sqrt{(-\tfrac{1}{3}k)}\}] = [3^2 \,.\, D^3] = J_1^2 J_2^2 D_1^3 D_2^3,$$

where J_1, J_2, D_1, D_2 are ideals in the field $R\sqrt{\{(-\tfrac{1}{3}k)\}}$, $J_1 J_2 = [3]$, and $D_1 D_2 = [D]$.

Let h be defined as before. Then J_1^h, J_2^h, D_1^h, D_2^h are principal ideals, and, since the fundamental units are ± 1, we have, with arbitrary choice of signs,

$$\{A \pm B\sqrt{(-\tfrac{1}{3}k)}\}^h = \{\tfrac{1}{2}\gamma \pm \tfrac{1}{2}\delta\sqrt{(-\tfrac{1}{3}k)}\}\{\tfrac{1}{2}E \pm \tfrac{1}{2}F\sqrt{(-\tfrac{1}{3}k)}\}^3,$$

where γ, δ, E, F are integers, and the two integers in each pair of γ, δ and E, F are either both odd or both even, and also γ, δ satisfy

$$\tfrac{1}{4}(\gamma^2 + \tfrac{1}{3}k\delta^2) = 3^{2h}.$$

Since $-\frac{1}{3}k \equiv t^2 \pmod{3^{2h}}$ is soluble in t we have

$$\gamma^2 - t^2\delta^2 \equiv 0 \pmod{3^{2h}}.$$

We may suppose that t is defined by

$$\gamma - t\delta \equiv 0 \pmod{3^{2h}}.$$

Then we obtain either

$$(A+Bt)^h \equiv 0 \pmod{3^{2h}},$$

$$A+Bt \equiv 0 \pmod{9};$$

or

$$(A-Bt)^h \equiv 0 \pmod{3^{2h}},$$

$$A-Bt \equiv 0 \pmod{9}.$$

Hence $$(\pm 2Bt)^h \equiv (\delta t)(\tfrac{1}{2}E \pm \tfrac{1}{2}tF)^3 \pmod{9},$$

where all the $\pm$ signs may be chosen arbitrarily.

We may exclude the case $\frac{1}{2}E \pm \frac{1}{2}tF \equiv 0 \pmod 3$ since then $A \equiv B \equiv 0 \pmod 3$. Hence $(\frac{1}{2}E \pm \frac{1}{2}tF)^3 \equiv \pm 1 \pmod 9$, and then

$$\delta \equiv \pm 2^h t^{h-1} \pmod 9.$$

If $h = 3n+1$, we have $\delta \equiv \pm 2 \pmod 9$.

If $h = 3n-1$, we have $\delta \equiv \pm 4(\frac{1}{3}k)^2 \pmod 9$.

Therefore there are no solutions in this case, if

(i) $\delta \not\equiv \pm 2 \pmod 9$, when $h = 3n+1$;

(ii) $\delta \not\equiv \pm 4(\frac{1}{3}k)^2 \pmod 9$, when $h = 3n-1$.

REFERENCES

1. L. J. Mordell, 'On some Diophantine equations $y^2 = x^3 + k$ with no rational solutions': *Archiv for math. og natur., B.I.L.*, Nr 6, 1947.
2. R. Fueter, 'Über kubische diophantische Gleichungen': *Comm. Math. Helvetici*, 2 (1930), 69–89.
3. O. Brunner, 'Lösungseigenschaften der kubischen diophantischen Gleichung': Dissertation (Zurich, 1933).
4. G. Billing, 'Beiträge zur arithmetischen Theorie ebener kubischer Kurven': *Nova Acta Reg. Soc. Sci. Upsal.* (4) 11 (i) (1938), 1–165. There is a table at the end of this paper giving a complete basis (in Poincaré's sense) of the rational solutions of $y^2 = x^3 + k$, where k is any integer numerically less than 25.

ON THE EQUATION $ax^2+by^2+cz^2 = 0$

By H. DAVENPORT (*London*) and
MARSHALL HALL (*Columbus, Ohio*)

[Received 19 March 1948]

LET a, b, c be integers, and suppose that (i) not all of a, b, c have the same sign, and none of them is zero, (ii) a, b, c are relatively prime in pairs. A famous theorem of Legendre states that, if the congruences

$$A^2 \equiv -bc \pmod{a}, \quad B^2 \equiv -ca \pmod{b}, \quad C^2 \equiv -ab \pmod{c} \tag{1}$$

are soluble, then the equation

$$ax^2+by^2+cz^2 = 0 \tag{2}$$

has a solution in integers x, y, z, not all zero. If the additional hypothesis is made that a, b, c are square-free, then the above condition is necessary, as well as sufficient, for the solubility of (2).

In this note, we give a simple proof of Legendre's theorem by using the methods of the geometry of numbers. The idea of the proof is the natural one of determining x, y, z so that $ax^2+by^2+cz^2$ is both divisible by abc and numerically less than $|abc|$.

The points (x, y, z), where x, y, z are any integers satisfying

$$Ay \equiv cz \pmod{a}, \quad Bz \equiv ax \pmod{b}, \quad Cx \equiv by \pmod{c}, \tag{3}$$

form a lattice in three-dimensional space, since the sum or difference of two such points is again such a point. Three particular lattice points are

$$(bc, 0, 0), \qquad (x_1, a, 0), \qquad (x_2, y_2, 1)$$

for appropriate values of x_1, x_2, y_2. Moreover, every solution of (3) is expressible in the form

$$(x, y, z) = \lambda(bc, 0, 0)+\mu(x_1, a, 0)+\nu(x_2, y_2, 1)$$

with integral λ, μ, ν, as one easily sees by determining ν, μ, λ successively. Since the determinant of the coordinates of the three points is abc, it follows that the determinant of the lattice is $|abc|$. All points of the lattice satisfy

$$ax^2+by^2+cz^2 \equiv 0 \pmod{abc}. \tag{4}$$

For $$c(by^2+cz^2) \equiv -A^2y^2+c^2z^2 \equiv 0 \pmod{a}$$

by (1) and (3), whence (4) follows as a congruence to modulus a, and similarly to moduli b and c.

Suppose, as we may after (i), that $a > 0$, $b > 0$, $c = -c' < 0$. The real linear transformation

$$x = X\sqrt{(bc')}, \qquad y = Y\sqrt{(ac')}, \qquad z = Z\sqrt{(ab)} \tag{5}$$

transforms the lattice in (x, y, z)-space into a lattice in (X, Y, Z)-space whose determinant is 1. For every point of this new lattice,

$$X^2+Y^2-Z^2 = (ax^2+by^2+cz^2)/(abc'),$$

and so is an integer. To complete the proof of Legendre's theorem, it will suffice to prove that every lattice of determinant 1 contains a point, other than the origin O, which satisfies*

$$|X^2+Y^2-Z^2| < 1. \tag{6}$$

Let M denote the lower bound of $|X^2+Y^2-Z^2|$ for all points of the lattice other than O. If $M < 1$, the desired conclusion holds, so we may suppose that $M \geqslant 1$. There exists a lattice point (X_1, Y_1, Z_1) for which

$$X_1^2+Y_1^2-Z_1^2 = \pm M_1, \tag{7}$$

where M_1 either equals M, or differs from M by an arbitrarily small amount.

Case 1. Suppose first that the lower sign holds in (7). We can find a real linear transformation which will leave $X^2+Y^2-Z^2$ invariant, and transform (X_1, Y_1, Z_1) into $(0, 0, \sqrt{M_1})$. For we can first transform (X_1, Y_1, Z_1) into $(0, \sqrt{(X_1^2+Y_1^2)}, Z_1)$ by a transformation of X, Y only, leaving X^2+Y^2 invariant, and then transform $(0, \sqrt{(X_1^2+Y_1^2)}, Z_1)$ into $(0, 0, \sqrt{M_1})$ by a transformation of Y, Z only, leaving Y^2-Z^2 invariant.

We now have a lattice in (X, Y, Z)-space, of determinant 1, such that $(0, 0, \sqrt{M_1})$ is a lattice point, and such that

$$|X^2+Y^2-Z^2| \geqslant M \tag{8}$$

for every lattice point except O. The points $(X, Y, 0)$ obtained by projecting all lattice points on the plane $Z = 0$ form a two-dimensional lattice of determinant $1/\sqrt{M_1}$. We apply the well-known theorem that any plane lattice of determinant Δ contains a lattice point, other than O, in the circle $X^2+Y^2 \leqslant 2\Delta/\sqrt{3}$. Thus there is a

* In fact, by a theorem of Markoff, this is true with any number greater than $\sqrt{(2/3)}$ in place of 1 on the right of (6). For an elementary proof of Markoff's theorem, see Davenport, *J. of London Math. Soc.* 22 (1947), 96–9.

point of the three-dimensional lattice satisfying

$$X^2+Y^2 \leqslant 2/\sqrt{(3M_1)}. \tag{9}$$

We can suppose, without loss of generality, by subtracting a multiple of $(0, 0, \sqrt{M_1})$ and changing signs throughout if necessary, that

$$0 \leqslant Z \leqslant \tfrac{1}{2}\sqrt{M_1}. \tag{10}$$

Thus $$X^2+Y^2-Z^2 \geqslant -\tfrac{1}{4}M_1,$$

and, since M_1 is approximately M, (8) implies that

$$X^2+Y^2-Z^2 \geqslant M. \tag{11}$$

Another lattice point is $(X, Y, Z-\sqrt{M_1})$, and, by (9) and 10,

$$X^2+Y^2-(Z-\sqrt{M_1})^2 \leqslant 2/\sqrt{(3M_1)}-\tfrac{1}{4}M_1.$$

Since M_1 is approximately M, and $M \geqslant 1$, the number on the right is less than M. Hence (8) gives

$$X^2+Y^2-(Z-\sqrt{M_1})^2 \leqslant -M. \tag{12}$$

On subtraction, (11) and (12) imply

$$2Z\sqrt{M_1} \leqslant M_1-2M < 0,$$

contrary to (10).

Case 2. Suppose now that the upper sign holds in (7). We can suppose also that there are no values of $X^2+Y^2-Z^2$ arbitrarily near to $-M$, since that possibility has been settled in Case 1. We can find (in the same way as before) a real linear transformation which leaves $X^2+Y^2-Z^2$ invariant, and transforms (X_1, Y_1, Z_1) into $(\sqrt{M_1}, 0, 0)$. Thus we have a lattice in (X, Y, Z)-space of determinant 1 such that $(\sqrt{M_1}, 0, 0)$ is a lattice point, and such that (8) holds for all lattice points except O. The points $(0, Y, Z)$, obtained by projecting the lattice points on the plane $X = 0$, form a lattice of determinant $1/\sqrt{M_1}$. We use the well-known theorem that a plane lattice of determinant Δ has a point, other than O, in the rectangle

$$|Y| < \lambda, \qquad |Z| < \mu,$$

if λ, μ are positive numbers satisfying $\lambda\mu > \Delta$. Hence there is a point of the three-dimensional lattice satisfying

$$|Y| < (M-\tfrac{1}{4}M_1)^{\frac{1}{2}}, \qquad |Z| < M+M_1.$$

For, since M_1 is approximately M and $M \geqslant 1$, the product of the numbers on the right is greater than $1/\sqrt{M_1}$. We can suppose without loss of generality that this lattice point has

$$0 \leqslant X \leqslant \tfrac{1}{2}\sqrt{M_1}. \tag{13}$$

Now $$X^2+Y^2-Z^2 < \tfrac{1}{4}M_1+(M-\tfrac{1}{4}M_1) = M;$$

hence (8) implies $$X^2+Y^2-Z^2 \leqslant -M. \tag{14}$$

Another lattice point is $(X+\sqrt{M_1}, Y, Z)$, and

$$(X+\sqrt{M_1})^2+Y^2-Z^2 > M_1-(M+M_1) = -M.$$

Hence (8) gives $$(X+\sqrt{M_1})^2+Y^2-Z^2 \geqslant M. \tag{15}$$

By (14) and (15), $$2X\sqrt{M_1} \geqslant 2M-M_1.$$

So, by (13), X is nearly $\frac{1}{2}\sqrt{M}$. Also

$$M-(X+\sqrt{M_1})^2 \leqslant Y^2-Z^2 \leqslant -M-X^2.$$

So Y^2-Z^2 is nearly $-\frac{5}{4}M$. But then $X^2+Y^2-Z^2$ is nearly $-M$, which contradicts the hypothesis made earlier.

A proud tradition . . .

Since 1700, whenever professors or research students in Cambridge have required some obscure mathematical treatise or an important foreign scientific book, they have ordered it from DEIGHTON BELL, and rarely have they been disappointed.

Today, even in this difficult post-war period, we are maintaining our high standard of service and are supplying scientific and technical books to all parts of the world.

Why not avail yourself of this same efficient service and the next time you are having difficulty in finding the book you require, write to

DEIGHTON, BELL & CO., LTD.

13 Trinity Street · Cambridge

N.B.—Our Bulletin of Foreign Books No. 6 is now ready and will be sent on request.

Ready shortly

DIVERGENT SERIES

By the late G. H. HARDY. *Probable price* 30*s. net*

This book by the late Professor Hardy contains the matter of his courses of lectures on Divergent Series, revised and enlarged in form suitable for publication. The first two chapters are mainly historical. The remaining eleven develop the general theorems and special methods whereby it has been found possible to manipulate such series in a rigorous manner. They deal with general methods of summation, arithmetic means, Tauberian theorems for power series, the methods of Euler and Borel, multiplication of series, Hausdorff means, Weiner's Tauberian theorems, and the Euler-Maclaurin sum formula.

By the same Author, with E. M. Wright:

AN INTRODUCTION TO THE THEORY OF NUMBERS. 25*s. net*

THE MATHEMATICS OF GREAT AMATEURS

By J. L. COOLIDGE. *Probable price* 21*s. net*

In this book Professor Coolidge describes the contributions made to the development of mathematics by men whom he has called amateurs (i.e. who are known chiefly for some other activity). He discusses the significance of their achievements in relation to both the body of mathematical knowledge existing at the time they were working and also to subsequent developments of the subject. The book should interest all who like reading about the history of mathematics and should be useful to university and school teachers. A feature of the book is the large number of text-figures which are reproductions of those actually used by these mathematicians when first publishing their work. The 'amateurs' whose work is discussed are: Plato; Omar Khayyám; Pietro dei Franceschi; Leonardo da Vinci; Albrecht Duerer; John Napier, Baron Merchiston; Blaise Pascal; Antoine Arnauld; Jan De Witt; J. Heinrich Hudde; William, Viscount Brouncker; Guillaume l'Hospital, Marquis de Sainte-Mesme; Buffon; Denis Diderot; William George Horner; Bernhard Bolzano.

By the same Author:

A HISTORY OF THE CONIC SECTIONS AND QUADRIC SURFACES. 25*s. net*
THE GEOMETRY OF THE COMPLEX DOMAIN. 25*s. net*
A HISTORY OF GEOMETRICAL METHODS. 30*s. net*
THE ELEMENTS OF NON-EUCLIDEAN GEOMETRY. 21*s. net*

OXFORD UNIVERSITY PRESS